

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 18.1432 Seconds
(without alignments)
636.214 Million cell updates/sec

Title: US-10-036-444-4

Perfect score: 632

Sequence: 1 LWSQPPPIRTLEGSSAFPL.....TGNGTRLVVEKEHPQLGAGT 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283365

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	18.4	132	2 A24402	T-cell receptor al
2	114.5	18.1	136	2 B45893	T-cell receptor al
3	107.5	17.0	138	2 C27577	T-cell receptor al
4	105	16.6	131	2 F45893	T-cell receptor al
5	103	16.3	129	2 A42692	T-cell receptor al
6	102.5	16.2	110	2 B24092	T-cell receptor al
7	101	16.0	120	2 I54487	T-cell receptor al
8	99.5	15.7	131	2 E24092	T-cell receptor al
9	99.5	15.7	131	2 D24092	T-cell receptor al
10	99	15.7	132	1 RWSAV	T-cell receptor al
11	98.5	15.6	130	2 A31211	T-cell receptor al
12	98.5	15.6	146	2 S26408	T-cell receptor be
13	97	15.3	139	2 S36325	T-cell receptor de
14	95.5	15.1	271	2 B47112	myelin/oligodendro
15	95.5	15.1	218	2 A53268	T-cell receptor al
16	94	14.9	526	2 S70587	butyrophilin precu
17	93.5	14.8	247	2 A57117	myelin/oligodendro
18	93	14.7	110	2 A24092	T-cell receptor al
19	93	14.7	117	2 I68824	T-cell receptor al
20	93	14.7	1694	2 S50065	sialoadhesin - mou
21	92.5	14.6	132	2 D45893	T-cell receptor al
22	92.5	14.6	139	2 S36302	T-cell receptor de
23	92.5	14.6	267	1 RWSVC8	T-cell receptor al
24	91.5	14.5	107	2 S60590	Ig heavy chain var
25	90	14.2	132	2 A27632	T-cell receptor al
26	89	14.1	134	2 A45893	T-cell receptor al
27	88.5	14.0	137	2 C45893	T-cell receptor al
28	88.5	14.0	223	2 A29063	cytotoxic T-lympho
29	88.5	14.0	264	2 F27579	T-cell receptor al

30	88.5	14.0	1197	2 T30581	neural cell adhesi
31	88	13.9	131	2 D24402	T-cell receptor al
32	88	13.9	226	2 A45477	membrane-bound imm
33	88	13.9	946	1 A47299	ror-related recept
34	87.5	13.8	111	2 JH0333	T-cell receptor al
35	87.5	13.8	132	2 S23374	T-cell receptor al
36	87.5	13.8	246	2 A47712	myelin/oligodendro
37	87.5	13.8	247	2 S58394	T-cell receptor al
38	87	13.8	108	2 JH0342	T-cell receptor al
39	86.5	13.7	95	2 G30603	T-cell receptor al
40	86.5	13.7	129	2 S03478	hypothetical prote
41	86.5	13.7	874	2 T29548	sax-3 protein - Ca
42	86.5	13.7	1273	2 T42405	Ig lambda chain pr
43	86	13.6	140	2 PH0134	T-cell receptor al
44	85.5	13.5	134	2 C29774	T-cell receptor al
45	85.5	13.5	223	2 T09536	cytotoxic T-lympho

ALIGNMENTS

RESULT 1

A24402
T-cell receptor alpha chain precursor V region (CS) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 30-May-1997
C/Accession: A24402
R/Becker, D.M.; Patten, P.; Chien, Y.; Yokota, T.; Eshhar, Z.; Giedlin, M.; Gascoigne, J.
Nature 317, 430-434, 1985
A/Title: Variability and repertoire size of T-cell receptor V-alpha gene segments.
A/Reference number: A93368, UID:86014379; PMID:2995827
A/Accession: A24402
A/Molecule type: mRNA
A/Residues: 1-132 <BEC>
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: T-cell receptor

Query Match 18.4%; Score 116; DB 2; Length 132;
Best Local Similarity 33.3%; Pred. No. 0.00019;
Matches 41; Conservative 16; Mismatches 42; Indels 24; Gaps 8;

Qy	3	VSQPPPIRTLEGSSAFPLCSFNASQRLAIGSVTVFVPGKEVR-----NGTPE 54
Db	23	VQSPESLIVPEGAMVSLNCSFSDS----ASQSIWYQQH--PGKPKALISIFSNK 76
Qy	55	FRQLAPLASSRPLHQRHLEIRVDHDAISIVCRVEVLGLVGT---GNGTRLVVEK 111
Db	77	EGRLTVYLNRLASLH---VSLHIKQSPSDSAVYLCAVRRSGANTGKLTFGHGTILRV-- 130
Qy	112	EHP 114
Db	131	-HP 132

RESULT 2

B45893
T-cell receptor alpha chain precursor V region (BTA29) - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C/Accession: B45893
R/Shiguro, N.; Tanaka, A.; Shinagawa, M.
Immunogenetics 31, 57-60, 1990
A/Title: Sequence analysis of bovine T-cell receptor alpha chain.
A/Reference number: A45893, UID:90129157; PMID:2137108
A/Accession: B45893
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-136 <ISH>
A/Cross-references: GB:D90011; NID:G217610; PIDN:BA414061.1; PID:G217611
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: T-cell receptor

Query Match 18.1%; Score 114.5; DB 2; Length 136;

Best Local Similarity 29.5%; Pred. No. 0.00028;
Matches 36; Conservative 19; Mismatches 39; Indels 29; Gaps 7;
QY 4 SQPPEIRTLGSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKEVRNGTPER-----55
Db 27 AEPASLPVPEGAASLGCTVSDNSLY----FWYRQ--YPGK-----GPEFLQVYANN 75
QY 56 ---RGLAPLASSRFLHDHQAELHIRDVRGHDASIVVCRVEVLGLGVGT-----GNGTRLV 108
Db 76 NKEGKFT--AQNKNKH-VSLIRDSEPSDATYLCAVDNISTAGTKLTFEGGTRLI 132
QY 109 VE 110
Db 133 VK 134

RESULT 3
T-cell receptor alpha chain V region (5/10-20D) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-May-1997
C:Accession: C27577; C27557
R: Iwamoto, A.; Ohashi, P.S.; Pիրcher, H.; Walker, C.L.; Michalopoulos, E.E.; Rupp, F.; H
J. Exp. Med. 165, 591-600, 1987
A:Title: T cell receptor variable gene usage in a specific cytotoxic T cell response. P
A:Reference number: A27557; MUID:87139812; PMID:3453320
A:Accession: C27577
A:Molecule type: mRNA
A:Residues: 1-138 <IWA>
A:Accession: C27557
A:Molecule type: mRNA
A:Residues: 22-138 <TW2>
A:Cross-references: EMBL:X05733
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 17.0%; Score 107.5; DB 2; Length 138;
Best Local Similarity 31.9%; Pred. No. 0.0014;
Matches 38; Conservative 17; Mismatches 45; Indels 19; Gaps 7;
QY 3 VSQPEIRTLGSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKEVR-----NGTPE 54
Db 24 VQSPESLIIVPEGAMTSLNCTFSDASQY----FAWYRQH--SGKAPKALMIFSNGEKE 77
QY 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDASIVVCRVEVL-GLGVGTGNGTRLVVEKE 112
Db 78 -EGFTTHNKASLH---FSLHIRDSPDSALYLCAVSTGCKYKVFSGGIRLLVSPD 132

RESULT 4
T-cell receptor alpha chain precursor V region (BTA25) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C:Accession: F45893
R: Ishiguro, N.; Tanaka, A.; Shinagawa, M.
Immunogenetics 31, 57-60, 1990
A:Title: Sequence analysis of bovine T-cell receptor alpha chain.
A:Reference number: A45893; MUID:90129157; PMID:2137108
A:Accession: F45893
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-131 <ISH>
A:Cross-references: GB:D90015; NID:G217618; PID:BAAL4065.1; PID:G217619
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:36-109/Domain: immunoglobulin homology <IMM>

Query Match 16.6%; Score 105; DB 2; Length 131;
Best Local Similarity 30.3%; Pred. No. 0.0024;
Matches 36; Conservative 23; Mismatches 34; Indels 26; Gaps 8;
QY 3 VSQPEIRTL-EGSSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKE-----VRNGTPEF 55

Db 24 VEQSPSVLSQEGANSILRCNFSDT-----VDSVQWFQON--PGALTTLTFFIASGTHK- 75
QY 56 RGLAPLASSRFLHDHQAELHIRDVRGHDASIVVCRVEVLGLGVGT-----GNGTRLV 109
Db 76 NERMSSTVNSK---ERYSTLHTASQLEDAATYLCVD---LGSNGRLVFGKGRILAV 128
RESULT 5
A42692
T-cell receptor alpha chain (lysozyme peptide specific) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999
C:Accession: A42692
R: Kobori, J.A.; Hood, L.; Shastri, N.
Proc. Natl. Acad. Sci. U.S.A. 91, 2940-2944, 1992
A:Title: Structure-function relationship among T-cell receptors specific for lysozyme p
A:Reference number: A42692; MUID:92212944; PMID:1313573
A:Accession: A42692
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-129 <KOB>
A:Cross-references: GB:M87844; NID:G201804; PID:AAAS1236.1; PID:G554375
A:Experimental source: T-cell hybrid BO4H.9.1
A:Note: sequence extracted from NCBI backbone (NCBIN:92850, NCBI:P:92851)
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:36-109/Domain: immunoglobulin homology <IMM>

Query Match 16.3%; Score 103; DB 2; Length 129;
Best Local Similarity 29.8%; Pred. No. 0.0037;
Matches 39; Conservative 14; Mismatches 34; Indels 44; Gaps 8;
QY 2 WVS-----QPPEIRTLGSAFLPCSFNASQGRLAIGSVTWFRDE-----VWPG 45
Db 18 WVSQDKVKQSPSALSQEGTNSALRCNFS-----IAATTVQWFLQNGRSLINLFYLV- 71
QY 46 KEVNGTPPEPRGLAPLASSRFLHDHQAELHIRDVRGHDASIVVCRVEVLGLGVGTGN-- 103
Db 72 -----GTYE-NGRLKSAFDSK---ESYSTLHIDAQLEDSGTFFCAAE-----DTGNYK 116
QY 104 -----GTRLVV 109
Db 117 YVFGAGTRLKV 127

RESULT 6
B24092
T-cell receptor alpha chain V region (5C.C7) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 21-Jan-2000
C:Accession: B24092
R: Fink, P.J.; Matlis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A:Title: Correlations between T-cell specificity and the structure of the antigen recep
A:Reference number: A93380; MUID:86230843; PMID:3012351
A:Accession: B24092
A:Molecule type: mRNA
A:Residues: 1-110 <FTN>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:16-89/Domain: immunoglobulin homology <IMM>

Query Match 16.2%; Score 102.5; DB 2; Length 110;
Best Local Similarity 31.9%; Pred. No. 0.0035;
Matches 36; Conservative 16; Mismatches 46; Indels 15; Gaps 6;
QY 3 VSQPEIRTL-EGSSAFLPCSFNASQGRLAIGSVTNF-----RDEVVPGKEVRNGTPEFRG 57
Db 4 VEQSPSALSLEHGSGSALRCNFTT-----MRVQWFRKNSRGLINLFLYASGTKE-NG 57
QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIVVCRVEVLGLG-VGTGNGTRLVV 109

Db 58 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAEASNTKVVFGTGRLOV 107

RESULT 7

T-cell receptor alpha chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C/Accession: I54487; I58823
R:Nakajima, P.B.; Di Vincenzo, J.P.; Jameson, S.C.; Gascoigne, N.R.J.
Immunogenetics 35, 190-198, 1992
A/RTitle: Chromosome 14 in B10.A (18R) mice is recombinant and includes Tcra-V alleles.
A/Reference number: I54487; MUID:92165347; PMID:1371499

A/Accession: I54487
A/Status: preliminary; translated from GB/EMBL/DBBJ
A/Molecule type: mRNA
A/Residues: 1-120 <RES>
A/Cross-references: GB:M55634; NID:gl99677; PIDN:AAA39701.1; PID:gl99678
A/Accession: I58823
A/Status: preliminary; translated from GB/EMBL/DBBJ
A/Molecule type: mRNA
A/Residues: 104-120 <RES2>
A/Cross-references: GB:M55634; NID:gl99677; PIDN:AAA39702.1; PID:gl99679
A/Note: J-alpha TA61
C/Genetics:

A/Gene: MHC-V-alpha-11; J-alpha-TA61
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: transmembrane protein
F:26-100/Domain: immunoglobulin homology <IMM>

Query Match 16.0%; Score 101; DB 2; Length 120;
Best Local Similarity 31.6%; Pred. No. 0.0054;
Matches 36; Conservative 17; Mismatches 47; Indels 14; Gaps 6;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFRD---EVVPGKEVRNGTPEFRG 57
DB 14 VEQSPALSLSHEGTDSALRCNFTTT-----MRVQWFRNRSRGLSLISFLYASGTYKE-NG 57

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGT-ONGTRLVVE 110
DB 68 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAADTNTGKLTFGDGTVLTKV 119

RESULT 8

T-cell receptor alpha chain precursor V region (4.C3) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 21-Jan-2000
C/Accession: E24092
R:Fink, P.J.; Matis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A/Title: Correlations between T-cell specificity and the structure of the antigen receptor
A/Reference number: A93380; MUID:86230843; PMID:3012351

A/Accession: E24092
A/Molecule type: mRNA
A/Residues: 1-131 <FIN>
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: T-cell receptor
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-131/Product: T-cell receptor alpha chain V region 4.C3 #status predicted <MAT>
F:36-109/Domain: immunoglobulin homology <IMM>

Query Match 15.7%; Score 99.5; DB 2; Length 131;
Best Local Similarity 29.3%; Pred. No. 0.0085;
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRVQWFRNRSRGLSLINFLYASGTYKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLG 117
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----STSSGQKLV-----FG 121

QY 118 AGT 120
DB 122 QGT 124

RESULT 9

D24092
T-cell receptor alpha chain precursor V region (B10) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Nov-1988 #sequence_revision 01-Dec-2000 #text_change 11-Jan-2002
C/Accession: D24092; S03507
R:Fink, P.J.; Matis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A/Title: Correlations between T-cell specificity and the structure of the antigen receptor
A/Reference number: A93380; MUID:86230843; PMID:3012351

A/Accession: D24092
A/Molecule type: mRNA
A/Residues: 1-131 <FIN>
R:Winoto, A.; Mjolsness, S.; Hood, L.
Nature 316, 832-836, 1985
A/Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.
A/Reference number: S03503; MUID:85256332; PMID:2993908

A/Accession: S03507
A/Molecule type: DNA
A/Residues: 111-130 <WIN>
A/Cross-references: EMBL:X03057; NID:954519; PIDN:CAA26864.1; PID:91334131
A/Note: this sequence was determined from the germline gene
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: T-cell receptor
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-131/Product: T-cell receptor alpha chain V region B10 #status predicted <MAT>
F:36-109/Domain: immunoglobulin homology <IMM>

Query Match 15.7%; Score 99.5; DB 2; Length 131;
Best Local Similarity 29.3%; Pred. No. 0.0085;
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRVQWFRNRSRGLSLINFLYASGTYKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLG 117
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----ATSSGQKLV-----FG 121

QY 118 AGT 120
DB 122 QGT 124

RESULT 10

RWMSAV
T-cell receptor alpha chain precursor V region (2B4) - mouse
C:Species: Mus musculus (house mouse)
C/Accession: A02015
R:Chien, Y.; Becker, D.M.; Lindsten, T.; Okamura, M.; Cohen, D.I.; Davis, M.M.
Nature 312, 31-35, 1984
A/Title: A third type of murine T-cell receptor gene.
A/Reference number: A93344; MUID:85036634; PMID:6548551
A/Accession: A02015
A/Molecule type: mRNA
A/Residues: 1-132 <CHI>
A/Experimental source: hybridoma 2B4, clone Tt11
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: glycoprotein; heterotetramer; receptor; T-cell
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-132/Product: T-cell receptor alpha chain V region (2B4) #status predicted <MAT>
F:21-113/Region: V segment
F:114-117/Region: D segment
F:118-132/Region: J segment
F:42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.7%; Score 99.5; DB 2; Length 131;
Best Local Similarity 29.3%; Pred. No. 0.0085;
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRVQWFRNRSRGLSLINFLYASGTYKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLG 117
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----STSSGQKLV-----FG 121

Query Match 15.78; Score 99; DB 1; Length 132;
 Best Local Similarity 29.48; Pred. No. 0.0096;
 Matches 37; Conservative 17; Mismatches 42; Indels 30; Gaps 8;

QY 3 VSQPE-IRTEGSAFLPCSFNASQGRLAIGSVTWFRDEVPGKEVR-----NGTPE 54
 DB 24 VQSPESLIVPEGARTSLNCTFSDASQY----FWYRQH--SGKAPKALMSIFSNGEKE 77

QY 55 FRGLAPLASSRFLHDHQAELHVRDVRGHDAIYVCRVEVLGVGTGNGTRLVVEKEHP 114
 DB 78 -EGFTTHLAKASLH--FSLHSDQPSDSALYCAVLTYG-----GSGNKLI----- 122

QY 115 QLCAGT 120
 DB 123 -FGTGT 127

RESULT 11
 A31211
 T-cell receptor alpha chain precursor V region (V11.3-J39) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999
 C:Accession: A31211
 R:Malissen, M.; Trucy, J.; Letourneur, F.; Rebai, N.; Dunn, D.E.; Fitch, F.W.; Hood, L.; Cell 55, 49-59, 1988
 A:Title: A T cell clone expresses two T cell receptor alpha genes but uses one alpha-beta
 A:Reference number: A94656; MUID:89003051; PMID:3262424
 A:Accession: A31211
 A:Molecule type: DNA
 A:Residues: 1-130 <MAL>
 A:Cross-references: GB:M22603; NID:G340716; PIDN:AAA63777.1; PID:G710556
 C:Genetics:
 A:Insertions: 18/1
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor
 F:36-109/Domain: immunoglobulin homology <IMW>

Query Match 15.68; Score 98.5; DB 2; Length 130;
 Best Local Similarity 30.68; Pred. No. 0.011;
 Matches 38; Conservative 15; Mismatches 42; Indels 29; Gaps 8;

QY 2 WVS-----OPPEIRTEGSAFLPCSFNASQGRLAIGSVTWFRDEVPGKEVRN----- 50
 DB 18 WVSQKVKQSPSALSLOGINSALRCNFS-----IAATTQWFLQN--PRGSLMNLFLV 70

QY 51 -GTPEFRGLAPLASSRFLHDHQAELHVRDVRGHDAIYVCRVEVLGVGT-----GNGT 105
 DB 71 PGTKE-NGRLLKSAFDSK---ESYSTLHDAQLSDSGTYFCAAG--GNSGTYYORFGTGT 124

QY 106 RLVV 109
 DB 125 KLVV 128

RESULT 12
 S26408
 T-cell receptor beta chain V region (5.5) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
 C:Accession: S26408
 R:Bowman, S.J.; Lanchbury, J.S.
 submitted to the EMBL Data Library, September 1992
 A:Description: T cell receptor beta chain sequences from patients with rheumatoid arthritis
 A:Reference number: S26408
 A:Accession: S26408
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-146 <BOH>
 A:Cross-references: EMBL:X68527; NID:G36172; PIDN:CAA48540.1; PID:G36173
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor
 F:35-112/Domain: immunoglobulin homology <IMW>

Query Match 15.68; Score 98.5; DB 2; Length 146;
 Best Local Similarity 26.08; Pred. No. 0.012;
 Matches 33; Conservative 23; Mismatches 44; Indels 27; Gaps 6;

QY 3 VSQPE-IRTEGSAFLPCSFNASQGRLAIGSVTWFRDEVPG-----KEVRNGT 52
 DB 23 VTQSPETHLTKT-RQQVTLRCSSQSGH-----NTVSNYQALGQGPQFIQYVREENG- 75

QY 53 PEFRGLAPLASSRFLHDHQAELHVRDVRGHDAIYVCRVEVLGVGT-----GNGT 105
 DB 76 ---RGNFPFRSGLOQFPNYSSELNVNALELDDLSALYLCASSFXGLGULPSRGYEQYFGPGT 132

QY 106 RLVIKE 112
 DB 133 RLTVTD 139

RESULT 13
 S36325
 T-cell receptor delta chain precursor - sheep (fragment)
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jan-2000
 C:Accession: S36325; S23040
 R:Hein, W.R.; Dudler, L.; EXHO J. 12, 715-724, 1993
 A:Title: Divergent evolution of T cell repertoires: extensive diversity and development
 A:Reference number: S36287; MUID:93178447; PMID:8440261
 A:Accession: S36325
 A:Molecule type: mRNA
 A>Status: preliminary; translation not shown
 A:Residues: 1-139 <HE>
 A:Cross-references: EMBL:Z12997; NID:G2258; PIDN:CAA78341.1; PID:G2259
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor
 F:33-113/Domain: immunoglobulin homology <IMW>

Query Match 15.38; Score 97; DB 2; Length 139;
 Best Local Similarity 30.68; Pred. No. 0.016;
 Matches 37; Conservative 10; Mismatches 52; Indels 22; Gaps 6;

QY 7 PEIRTEGSAFLPCSFNASQGRLAIGSVTWFR---DEVPGKEVRNGTPEFRGLAPLA 63
 DB 26 PEQVRASGREVTLOCTFTQTYSN---PDLYWYRKTDPDAVFQVLYRDNT---RSRDADFA 79

QY 64 SSRFLHDQA-----ELHVRDVRGHDAIYVCRVEVLGV-----GTGNGTRLVVEK 111
 DB 80 RGRFTVQHSVRSKTFHLVISSVRPDTATYTCALD-LQGIYDTDKLTGKGLRLIVEP 138

QY 112 E 112
 DB 139 E 139

RESULT 14
 B47712
 myelin/oligodendrocyte glycoprotein - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 21-Jan-1994 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
 C:Accession: B47712
 R:Pham-Dinh, D.; Mattei, M.G.; Nussbaum, J.L.; Rousset, G.; Pontarotti, P.; Roessel, N. Proc. Natl. Acad. Sci. U.S.A. 90, 7990-7994, 1993
 A:Title: Myelin/oligodendrocyte glycoprotein is a member of a subset of the immunoglobl
 A:Reference number: A47712; MUID:93376728; PMID:8367453
 A:Accession: B47712
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-218 <PHA>
 A:Experimental source: brain
 A:Note: sequence extracted from NCBI backbone (NCBIP:137804)
 C:Genetics:
 A:Gene: MOG
 C:Function:

Search completed: February 26, 2004, 12:14:50
Job time : 19.1432 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 10.1857 Seconds
(without alignments)
613.452 Million cell updates/sec

Title: US-10-036-444-4
Perfect score: 632
Sequence: 1 LWSQPPEIRTELGSSAFPLP.....TGNTRLVVEKEHPOLGAGT 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	15.7	132	1 TVA2_MOUSE	P01739 mus musculus
2	95.5	15.1	245	1 MOG_RAT	Q63345 rattus norv
3	94	14.9	526	1 BUTY_HUMAN	Q13410 homo sapien
4	93.5	14.8	246	1 MOG_MOUSE	Q61885 mus musculus
5	93	14.7	1694	1 SN_MOUSE	Q62230 mus musculus
6	88.5	14.0	223	1 CTL4_MOUSE	P09793 mus musculus
7	88.5	14.0	1197	1 CAM1_BRARE	Q90478 brachydanio
8	88	13.9	226	1 C79A_HUMAN	P11912 homo sapien
9	87.5	13.8	524	1 BUTY_MOUSE	Q62556 mus musculus
10	87.5	13.8	223	1 CTL4_HUMAN	P16410 homo sapien
11	87.5	13.8	246	1 MOG_BOVIN	P55803 bos taurus
12	87.5	13.8	247	1 MOG_HUMAN	Q16853 homo sapien
13	85.5	13.5	223	1 CTL4_PIG	Q9MYX7 sus scrofa
14	84.5	13.4	223	1 CTL4_RABIT	P42072 oryctolagus
15	83	13.1	246	1 MYPO_HETFR	P20938 heterodontu
16	82.5	13.1	403	1 CD33_MOUSE	Q63994 mus musculus
17	82	13.0	467	1 SIL5_MOUSE	Q91557 mus musculus
18	82	13.0	3707	1 PGBM_MOUSE	Q05793 mus musculus
19	81.5	12.9	646	1 MUI8_HUMAN	P43121 homo sapien
20	81	12.8	526	1 BUTY_BOVIN	P18892 bos taurus
21	79	12.5	551	1 SIL5_HUMAN	O15389 homo sapien
22	78.5	12.4	111	1 LV2D_HUMAN	P01710 homo sapien
23	78.5	12.4	463	1 SIL9_HUMAN	Q9Y336 homo sapien
24	77	12.2	134	1 TVB7_MOUSE	P06320 mus musculus
25	77	12.2	270	1 BASI_RABIT	Q28740 oryctolagus
26	77	12.2	385	1 AMPC_LYSLA	Q48743 lysobacter
27	76.5	12.1	365	1 CXAR_HUMAN	P78310 homo sapien
28	76	12.0	1051	1 PKT7_CHICK	Q91048 gallus gall
29	75	11.9	365	1 CXAR_MOUSE	P97792 mus musculus
30	74.5	11.8	111	1 LV1D_HUMAN	P01702 homo sapien
31	74.5	11.8	1452	1 PTPM_HUMAN	P28827 homo sapien
32	74	11.7	215	1 CIB2_HUMAN	O60939 homo sapien
33	74	11.7	977	1 KEMS_MOUSE	P09581 mus musculus

RESULT 1				
ID	TVA2_MOUSE	STANDARD;	PRT;	132 AA.
AC	P01739;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	T-cell receptor alpha chain V region 2B4 precursor.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (CLONE TT11).			
RX	MEDLINE=85036634; PubMed=6548551;			
RA	Chien Y., Becker D.M., Lindsten I., Okamura M., Cohen D.I.,			
RA	Davis M.M.;			
RT	"A third type of murine T-cell receptor gene.";			
RL	Nature 312:31-35(1984).			
DR	PIR; A02015; RMSAV.			
DR	HSSP; P01607; 1BEI.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF00047; Ig; 1.			
DR	SMART; SM00406; IG; 1.			
DR	PROSITE; PS00835; IG LIKE; 1.			
KW	T-cell; Receptor; Immunoglobulin domain; Glycoprotein; Signal.			
FT	SIGNAL 1 20			
FT	CHAIN 21 132			
FT	DOMAIN 21 113			
FT	DOMAIN 114 117			
FT	DOMAIN 118 132			
FT	CARBOHYD 42 42			
FT	NON TER 132 132			
SQ	SEQUENCE 132 AA; 14668 MW; CABBF6CF1DD3448B CRC64;			
Query Watch 15.7%; Score 99; DB 1; Length 132;				
Best Local Similarity 29.4%; Pred. No. 0.0025;				
Matches 37; Conservative 17; Mismatches 42; Indels 30; Gaps 8;				
QY	3 VSQPEE-IRTEGSSAFPLCSFNASQCRLAIGSVTFERDEVPGKEVR-----NGTPE 54			
Db	24 VQQPESLIVEGARTSLNCTFSASQY----FWYRQH--SGKAPKALMSIFSNGEKE 77			
QY	55 FGRGLAPLASRRFLHDHQAEHLHIRDVGHDSIVYVCRVEVLGLGVGTGNGTRLVVEKEHP 114			
Db	78 -EGRFTIHLNKASLH---FSLHIRDSPDSALYLCATVLYG-----GSGNKLI----- 122			
QY	115 QLQAGT 120			
Db	123 -FGTGT 127			
RESULT 2				
ID	MOG_RAT	STANDARD;	PRT;	245 AA.

AC Q63345;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myelin-oligodendrocyte glycoprotein precursor.
 GN MOG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=93085763; PubMed=1453482;
 RA Gardiner M.V., Amiguet P., Livingston C., Mathieu J.-M.;
 RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
 RT immunoglobulin superfamily.";
 RL J. Neurosci. Res. 33:177-187(1992).
 RN [2]
 RP SEQUENCE OF 28-245 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93376728; PubMed=8367453;
 RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,
 RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,
 RA Dautigny A.;
 RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
 RT immunoglobulin superfamily encoded within the major
 RT histocompatibility complex.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).
 RN [3]
 RP STRUCTURE BY NMR OF 62-82.
 RX MEDLINE=97354172; PubMed=9210466;
 RA Alabout-Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.;
 RT "A conformational study of the human and rat encephalitogenic myelin
 RT oligodendrocyte glycoprotein peptides 35-55.";
 RL Eur. J. Biochem. 246:59-70(1997).
 CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in
 CC completion and/or maintenance of the myelin sheath and in cell-
 CC cell communication.
 CC -!- SUBUNIT: May form homodimers.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
 CC localized on the surface of myelin and oligodendrocyte cytoplasmic
 CC membranes.
 CC -!- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
 CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
 CC ACTIVE MYELINATION.
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
 CC family.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
 CC with oligodendrocyte-myelin glycoprotein (OMG).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M99485; AAA41628.1; -;
 DR EMBL; L21995; AAF74786.1; -;
 DR PIR; B47712; B47712.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR Pfam; PF00047; IG_1.
 DR SMART; SM00406; IGv; 1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
 KW SIGNAL 1 27
 FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
 FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 156 176 POTENTIAL.

FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 209 229 POTENTIAL.
 FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 30 139 IG-LIKE.
 FT DISULFID 51 125 POTENTIAL.
 FT CARBOHYD 58 58 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 245 AA; 27881 MW; C97F8AD60D6A32B4 CRC64;
 Query Match 15.1%; Score 95.5; DB 1; Length 245;
 Best Local Similarity 36.2%; Pred. No. 0.011;
 Matches 34; Conservative 8; Mismatches 33; Indels 19; Gaps 6;
 QY 9 IRTLEGSSAFPLPCSFNASQRLAIG-SVTWFRDEVVPGKEV-----RNG-----TPEFR 56
 DB 39 IRLVGDDELPC--RISPGKNATGMEVGVYRS---PFSRVVHLYRNGKQDAEAPEYR 93
 QY 57 GLRLPLASSRFLHDHQAELHIRDVRGHDAIYVC 90
 DB 94 GRTELLAKES--IGGKVALRIQNVRFSDGGYTC 125
 RESULT 3
 BUTY_HUMAN
 ID BUTY_HUMAN STANDARD; PRT; 526 AA.
 AC Q13410;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).
 GN BTN1A1 OR BTN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Breast;
 RX MEDLINE=96201696; PubMed=8611614;
 RA Taylor M.R., Peterson J.A., Ceriani R.L., Couto J.R.;
 RT "Cloning and sequence analysis of human butyrophilin reveals a
 RT potential receptor function.";
 RL Biochim. Biophys. Acta 1306:1-4(1996).
 CC -!- FUNCTION: May function in the secretion of milk-fat droplets. It
 CC may act as a specific membrane-associated receptor for the
 CC association of cytoplasmic droplets with the apical plasma
 CC membrane (By similarity).
 CC -!- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
 CC family.
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U39576; AAC50489.1; -;
 DR PIR; S70587; S70587.
 DR Genew; HGNC:1135; BTN1A1.
 DR MIM; 601610; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR InterPro; IPR001870; B302.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003596; IG_v.
 DR InterPro; IPR006574; PRY.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00047; IG; 1.

```

DR Pfam: PF00622; SPRY; 1.
DR SMART: SM00406; IGV; 1.
DR SMART: SM00589; PRI; 1.
DR SMART: SM00449; SPRY; 1.
DR PROSITE: PS00835; IG_LIKE; 2.
KW Transmembrane; Glycoprotein; Immunoglobulin domain; Signal; Repeat.
FT SIGNAL 1 26
FT CHAIN 27 526
FT DOMAIN 27 242
FT TRANSMEM 243 259
FT DOMAIN 270 526
FT DOMAIN 27 138
FT DOMAIN 148 234
FT CARBOHYD 55 55
FT CARBOHYD 215 215
SQ SEQUENCE 526 AA; 59004 MW; E9EAC0CF8AP94D5 CRC64;

Query Match 14.9%; Score 94; DB 1; Length 526;
Best Local Similarity 39.6%; Pred. No. 0.037;
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;

QY 6 PPE-IRTLGSSAFPLPC--SFNASQGLAIGSVTFPRDVPVGKEV-RNG-----TPE 54
DB 34 PEPILAVGDEALPCRLSPNASAEHL---ELRFRKVKVPAVLVHRDGRQEAQWPE 90
QY 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVC 90
DB 91 YRGR-ATLVQDGIAGKRVV-LRIRGVRSDDGEYTC 124

RESULT 4
MOG MOUSE STANDARD; PRT; 246 AA.
AC Q61885; P70364; Q62003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=95130110; PubMed=7829100;
RA Daubas P., Pham-Dinh D., Dautigny A.;
RT "Structure and polymorphism of the mouse myelin/oligodendrocyte
RL glycoprotein gene.";
RL Genomics 23:36-41(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gardiner M.V., Matthieu J.M.;
RT "Murine and human MOG are highly conserved: cDNA analysis.";
RL Trans. Am. Soc. Neurochem. 24:234-234(1993).
RN [3]
RP SEQUENCE OF 29-246 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Roussel G.,
RA Pontarotti P., Roeckel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).
RN [4]
RP SEQUENCE OF 29-54.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=92218912; PubMed=1373175;
RA Aniguet P., Gardiner M.V., Zanetta J.-P., Matthieu J.-M.;
RT "Purification and partial structural and functional characterization
of mouse myelin/oligodendrocyte glycoprotein.";
J. Neurochem. 58:1676-1682(1992).
-!- FUNCTION: Minor component of the myelin sheath. May be involved in
completion and/or maintenance of the myelin sheath and in cell-
cell communication.
-!- SUBUNIT: May form homodimers.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
localized on the surface of myelin and oligodendrocyte cytoplasmic
membranes.
-!- DISEASE: Reduced concentrations of MOG are observed in jimpy and
quacking dysmyelinating mutant mice.
-!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTM/MOG
family.
-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
with oligodendrocyte-myelin glycoprotein (OMG).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; L29503; AAC42023.1; -.
EMBL; L29498; AAC42023.1; JOINED.
EMBL; L29500; AAC42023.1; JOINED.
EMBL; L29501; AAC42023.1; JOINED.
EMBL; L29499; AAC42023.1; JOINED.
EMBL; L29502; AAC42023.1; JOINED.
EMBL; U64572; AAB08096.1; -.
EMBL; L20942; AAB03180.1; -.
PIR; A55717; A55717.
DR MGD; MGI:97435; MOG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 246
FT DOMAIN 29 156
FT TRANSMEM 157 177
FT TRANSMEM 178 209
FT TRANSMEM 210 230
FT DOMAIN 231 246
FT DOMAIN 31 144
FT DISULFID 52 126
FT CARBOHYD 59 59
FT CONFLICT 21 21
FT CONFLICT 32 32
FT CONFLICT 95 95
FT CONFLICT 169 169
SQ SEQUENCE 246 AA; 28271 MW; 1F1A8A4A0D5CFB89 CRC64;

Query Match 14.8%; Score 93.5; DB 1; Length 246;
Best Local Similarity 35.1%; Pred. No. 0.018;
Matches 33; Conservative 9; Mismatches 33; Indels 19; Gaps 6;

QY 9 IRTLEGSAFLPCSFNASQGLAIG-SVTFPRDVPVGKEV-----RNG-----TPEFR 56
DB 40 IRLVGDGAELPC--RISPGKNATGMVGVWYS---PFSRVVHLRYRNGKDQDAEQAPEYR 94
QY 57 GLRLAPLASSRFLHDHQAELHIRDVRGHDASIYVC 90
DB 95 GRTELLKET--ISEKVTLRIONVRFSDGEYTC 126

RESULT 5
SN_MOUSE

```

ID SN MOUSE STANDARD; PRT; 1694 AA.
AC Q62230; O55216; Q62228; Q62229;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-
DE 1) (Sheep erythrocyte receptor) (SER).
GN SN OR SA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Macrophage;
RX MEDLINE=95009950; PubMed=7925291;
RA Crocker P.R.; Mucklow S.; Boukron V.; McWilliam A.; Willis A.C.;
RA Gordon S.; Wilson G.; Kelm S.; Bradfield P.;
RT "Sialoadhesin, a macrophage sialic acid binding receptor for
RT haemopoietic cells with 17 immunoglobulin-like domains.";
RL EMBO J. 13:4490-4503(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=98051930; PubMed=9383289;
RX Mucklow S.; Gordon S.; Crocker P.R.;
RA "Characterization of the mouse sialoadhesin gene, Sn.";
RL Mamm. Genome 8:934-937(1997).
RN [3]
RN CHARACTERIZATION.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=91266893; PubMed=2050106;
RA Crocker P.R.;
RT "Purification and properties of sialoadhesin, a sialic acid-binding
RT receptor of murine tissue macrophages.";
RL EMBO J. 10:1661-1669(1991).
RN [4]
RN SIALIC ACID BINDING.
RP MEDLINE=9517952; PubMed=7533044;
RX Kelm S.; Pelz A.; Schauer R.; Filbin M.T.; Tang S.; de Bellard M.E.;
RA Schnaar R.L.; Mahoney J.A.; Hartnell A.; Bradfield P.; Crocker P.R.;
RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new
RT family of sialic acid-dependent adhesion molecules of the
RT immunoglobulin superfamily.";
RL Curr. Biol. 4:965-972(1994).
RN [5]
RN BINDING TO SPN.
RP MEDLINE=21136329; PubMed=11238599;
RX van den Berg T.K.; Nath D.; Ziltener H.J.; Vestweber D.; Fukuda M.;
RA van Die I.; Crocker P.R.;
RT "CD43 functions as a T cell counterreceptor for the macrophage
RT adhesion receptor sialoadhesin (Siglec-1).";
RL J. Immunol. 166:3637-3640(2001).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 20-138.
RP MEDLINE=98325385; PubMed=9660955;
RA May A.P.; Robinson R.C.; Vinson M.; Crocker P.R.; Jones E.Y.;
RT "Crystal structure of the N-terminal domain of sialoadhesin in
RT complex with 3' sialyllactose at 1.85 A resolution.";
RL Mol. Cell 1:719-728(1998).
RN [7]
RN STRUCTURE OF 20-138 BY NMR, AND MUTAGENESIS OF TRP-21 AND ARG-116.
RP MEDLINE=99321481; PubMed=10393093;
RA Crocker P.R.; Vinson M.; Kelm S.; Drickamer K.;
RT "Molecular analysis of sialoside binding to sialoadhesin by NMR and
RT site-directed mutagenesis.";
RL Biochem. J. 341:355-361(1999).
CC -!- FUNCTION: Macrophage-restricted adhesion molecule that mediates
CC sialic-acid dependent binding to lymphocytes, including
CC granulocytes, monocytes, natural killer cells, B-cells and CD8 T-
CC cells (By similarity). Preferentially binds to alpha2,3-linked
CC sialic acid. Binds to SPN/CD43 on T-cells. May play a role in
CC hemopoiesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and
CC

soluble (isoforms 2 and 3).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=3;
Name=1;
IsoId=Q62230-1; Sequence=Displayed;
Name=2;
IsoId=Q62230-2; Sequence=VSP_002573; VSP_002574;
Name=3;
IsoId=Q62230-3; Sequence=VSP_002575; VSP_002576;
-!- TISSUE SPECIFICITY: Expressed by macrophages in various tissues.
Highest expression in spleen and lymph node with lower amounts in
lung, liver, bone marrow, heart and skin. No expression in thymus,
kidney, brain or small intestine.
-!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
(sialic acid binding Ig-like lectin) family.
-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; Z36293; CRA85290.1; -;
EMBL; Z36233; CRA85268.1; -;
EMBL; Z36234; CRA85269.1; -;
EMBL; U92842; AAB95641.1; -;
EMBL; U92833; AAB95641.1; JOINED.
EMBL; U92834; AAB95641.1; JOINED.
EMBL; U92836; AAB95641.1; JOINED.
EMBL; U92837; AAB95641.1; JOINED.
EMBL; U92838; AAB95641.1; JOINED.
EMBL; U92839; AAB95641.1; JOINED.
EMBL; U92840; AAB95641.1; JOINED.
EMBL; U92841; AAB95641.1; JOINED.
PIR; S50065; S50065.
PDB; 1QFO; 16-APR-99.
PDB; 1QFP; 16-APR-99.
MGD; MGI:99668; Sn.
DR GO; GO:0016021; C:integral to membrane; ISS.
DR GO; GO:0005529; F:sugar binding; ISS.
DR GO; GO:0016337; P:cell-cell adhesion; ISS.
DR GO; GO:0007160; P:cell-matrix adhesion; ISS.
DR GO; GO:0006954; P:inflammatory response; ISS.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_16.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS50835; IG_LIKE; 14.
DR PROSITE; PS00290; IG_MHC; 1.
KW Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW Immunoglobulin domain; Repeat; Alternative splicing; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1694 SIALOADHESIN.
FT DOMAIN 20 1638 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1639 1859 POTENTIAL.
FT DOMAIN 1660 1894 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 136 IG-LIKE V-TYPE.
FT DOMAIN 153 235 IG-LIKE C2-TYPE 1.
FT DOMAIN 239 321 IG-LIKE C2-TYPE 2.
FT DOMAIN 326 406 IG-LIKE C2-TYPE 3.
FT DOMAIN 416 508 IG-LIKE C2-TYPE 4.
FT DOMAIN 509 594 IG-LIKE C2-TYPE 5.
FT DOMAIN 602 701 IG-LIKE C2-TYPE 6.
FT DOMAIN 704 781 IG-LIKE C2-TYPE 7.
FT DOMAIN 795 890 IG-LIKE C2-TYPE 8.
FT DOMAIN 894 972 IG-LIKE C2-TYPE 9.
FT DOMAIN 979 1078 IG-LIKE C2-TYPE 10.
FT DOMAIN 1080 1160 IG-LIKE C2-TYPE 11.

```

1171 1263 IG-LIKE C2-TYPE 12.
1244 1336 IG-LIKE C2-TYPE 13.
1341 1438 IG-LIKE C2-TYPE 14.
1441 1519 IG-LIKE C2-TYPE 15.
1533 1626 IG-LIKE C2-TYPE 16.
SITE CELL ATTACHMENT SITE (POTENTIAL).
BY SIMILARITY.
FT DISULFID 36 166.
FT DISULFID 41 98.
FT DISULFID 160 218.
FT DISULFID 263 306.
FT DISULFID 347 391.
FT DISULFID 434 492.
FT DISULFID 532 576.
FT DISULFID 625 685.
FT DISULFID 725 770.
FT DISULFID 813 872.
FT DISULFID 911 955.
FT DISULFID 1000 1062.
FT DISULFID 1102 1144.
FT DISULFID 1188 1236.
FT DISULFID 1276 1319.
FT DISULFID 1362 1421.
FT DISULFID 1462 1508.
FT DISULFID 1551 1610.
FT CARBOHYD 159 159.
FT CARBOHYD 266 266.
FT CARBOHYD 299 299.
FT CARBOHYD 340 340.
FT CARBOHYD 500 500.
FT CARBOHYD 583 583.
FT CARBOHYD 693 693.
FT CARBOHYD 722 722.
FT CARBOHYD 737 737.
FT CARBOHYD 882 882.
FT CARBOHYD 1089 1089.
FT CARBOHYD 1099 1099.
FT CARBOHYD 1246 1246.
FT CARBOHYD 1459 1459.
FT CARBOHYD 1473 1473.
FT CARBOHYD 326 340.
FT VARSPLIC 341 1694.
FT VARSPLIC 1528 1598.
FT VARSPLIC 1599 1694.
FT MUTAGEN 21 21.
FT MUTAGEN 116 116.
FT MUTAGEN 116 116.
Query Match 14.7%; Score 93; DB 1; Length 1694;
Best Local Similarity 29.4%; Pred. No. 0.17;
Matches 32; Conservative 17; Mismatches 34; Indels 26; Gaps 8;
QY 2 W-VSQPEIRTLGSSAFLPCSFN-----ASGRLAIGSVTWFRDEVVPGKE---VRNG 51
Db 21 WGVSSPKNVGLSGSCLLICIFSVADVPVNSGITAL----WYD--YSGRRQVHSG 74
QY 52 TPE-----FGRGLAPLASSRFLHDHQ-AELHIRDVRGHSDASIVYCRREV 94
Db 75 DPKLVDRKFRGRAELMGNM-----DHKVCNLLKDLKPEDSGTNYNFRFEI 119
RESULT 6
CTLA4 MOUSE
ID CTL4 MOUSE STANDARD; PRT; 223 AA.
AC P09793; Q9QZ27;
DT 01-MAR-1989 (Rel. 10, Created)
101-MAR-1989 (Rel. 10, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-
associated antigen 4) (CTLA-4).
GN C14A4 OR CD152.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RX SEQUENCE FROM N.A.
MEDLINE=87258259; PubMed=3496540;
Brunet J.-F., Denizot F., Luciano M.-P., Roux-Dosseto M., Suzan M.,
Mattei M.-G., Golstein P.;
RA "A new member of the immunoglobulin superfamily -- CTLA-4.";
Nature 328:267-270(1987).
[2]
RX SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
STRAIN=129/SVJ;
MEDLINE=94425274; PubMed=10493833;
Ling V., Wu P.W., Finnerty H.F., Sharpe A.H., Gray G.S., Collins M.;
RA "Complete sequence determination of the mouse and human CTLA4 gene
loci: cross-species DNA sequence similarity beyond exon borders.";
Genomics 60:341-355(1999).
[3]
RX SEQUENCE OF 1-35 FROM N.A.
MEDLINE=91318145; PubMed=1713603;
Harper K., Balzano C., Rouvier E., Mattei M.-G., Luciano M.F.,
Golstein P.;
RA "CTLA-4 and CD28 activated lymphocyte molecules are closely related
in both mouse and human as to sequence, message expression, gene
structure, and chromosomal location.";
J. Immunol. 147:1037-1044(1991).
CC - FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
(CD80) AND B7-2 (CD86).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - TISSUE SPECIFICITY: Widely expressed with highest levels in
lymphoid tissues.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X05719; CAA29191.1; -.
EMBL; AF142145; AAF01489.1; -.
EMBL; M74362; AAA37489.1; -.
PIR; A29063; A29063.
HSSP; P16410; 1AH1.
MGP; MGI:88556; Ctl4.4.
InterPro; IPR008096; CTLA4.
InterPro; IPR007110; Ig-like.
Pfam; PF00047; Ig; 1.
PRINTS; PR01720; CTLANTIGEN4.
SMART; SM00409; Ig; 1.
PROSITE; PS50835; IG_LIKE; 1.
Immunoglobulin domain; T-cell, Transmembrane; Glycoprotein; Signal.
SIGNAL 1 35
CHAIN 36 223
DOMAIN 36 161
EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162 187
DOMAIN 188 223
CYTOPLASMIC (POTENTIAL).
FT DOMAIN 36 145
IG-LIKE V-TYPE.
FT DISULFID 58 129
BY SIMILARITY.
FT DISULFID 85 103
BY SIMILARITY.
FT CARBOHYD 108 108
N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 113 113
N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145
N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```
FT CONFLICT 182 182 S -> T (IN REF. 2).
SQ SEQUENCE 223 AA; 24993 MW; 5318FAAF416F4685 CRC64;

Query Match
Best Local Similarity 14.0%; Score 88.5; DB 1; Length 223;
Matches 33; Conservative 17; Mismatches 58; Indels 13; Gaps 5;

QY 3 VSQPPETRLTGGSAFLPCSFNAGQRLAIGSVTWFRD-----EVVPGKEVNRGTTPFR 56
Db 40 VTQPSVVLASHGVASFPCEYSBHTNDEV-RVTLRQNDQMTVCATFTKNTVGL 98
QY 57 GRLAPLASSRFLDHOAELHIRDVRGHDSIYVCRVEVL---GLGVGTGNGRLVVEKEH 113
Db 99 D-YPCSGGF-NESRVNLTQGLRAVDIGLYCKVELVPPPYFVGMGNGTQIVVIDPE 155
QY 114 P 114
Db 156 P 156

RESULT 7
CAML BRARE STANDARD; PRT; 1197 AA.
AC Q90478;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1.1 (N-CAM L1.1) (Fragment).
GN NADL1.1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Embryo;
RX MEDLINE=96155762; PubMed=8568941;
RA Tongiorgi E., Bernhardt R.R., Schachner M.;
RT "Zebrafish neurons express two L1-related molecules during early
axogenesis."
RL J. Neurosci. Res. 42:547-561(1995).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
development of the nervous system. Involved in neuron-neuron
adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
to axonin on neurons (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in postmitotic neurons in 16-36
hour embryos, including those in the brain, cranial ganglia and
otic and olfactory placodes, and in all classes of spinal
neurons.
CC -1- DEVELOPMENTAL STAGE: Onset of expression correlates with the
initiation of axogenesis in 16-36 hour embryos.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89204; CAA61490.1; .
CC F01; T30581; T30581.
CC HSP; P20241; 1CFB.
CC ZFIN; ZDB-GENE-980526-512; nadl1.1.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003962; FNIII_subd.
```

```
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00041; fn3; 5.
DR PRINTS; PR00047; Ig; 6.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00408; IGG2; 4.
DR PROSITE; PS50835; IG LIKE; 6.
KW Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
KW transmembrane; Repeat; Immunoglobulin domain.
FT DOMAIN 1 1054 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1055 1075 POTENTIAL.
FT DOMAIN 1076 1197 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 58 IG-LIKE C2-TYPE 1.
FT DOMAIN 69 160 IG-LIKE C2-TYPE 2.
FT DOMAIN 165 263 IG-LIKE C2-TYPE 3.
FT DOMAIN 268 355 IG-LIKE C2-TYPE 4.
FT DOMAIN 360 442 IG-LIKE C2-TYPE 5.
FT DOMAIN 451 541 IG-LIKE C2-TYPE 6.
FT DOMAIN 546 638 FIBRONECTIN TYPE-III 1.
FT DOMAIN 645 739 FIBRONECTIN TYPE-III 2.
FT DOMAIN 744 849 FIBRONECTIN TYPE-III 3.
FT DOMAIN 850 948 FIBRONECTIN TYPE-III 4.
FT DOMAIN 952 1029 FIBRONECTIN TYPE-III 5.
FT DISULFID 92 143 BY SIMILARITY.
FT DISULFID 199 247 BY SIMILARITY.
FT DISULFID 289 339 BY SIMILARITY.
FT DISULFID 383 432 BY SIMILARITY.
FT DISULFID 472 525 BY SIMILARITY.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1002 1002 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1197 AA; 132860 MW; 7CE1505EEFAC7B28 CRC64;

Query Match 14.0%; Score 88.5; DB 1; Length 1197;
Best Local Similarity 25.9%; Pred. No. 0.34; Mismatches 17; Gaps 4;
Matches 29; Conservative 14; Indels 52;

QY 3 VSQPPETRLTGGSAFLPCSFNAGQRLAIGSVTWFRDDEVVPGKEVNRGTTPFRGLAPL 62
Db 454 VGPQNHLVIRGSDAILHCKYTVDH-NLKSPTVQWKKD-----GHKITAST----- 498
QY 63 ASSRFLDHOAELHIRDVRGHDSIYVCRVEVLGLGVGTGNGRLVVEKEHP 114
Db 499 -SNDKTHEIGSLKVLVDQMEDMGISYCEVSTT-LDSDTASGVITVQDKPDP 548

RESULT 8
C79A HUMAN
ID C79A_HUMAN STANDARD; PRT; 226 AA.
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE B-cell antigen receptor complex associated protein alpha-chain
precursor (Ig-alpha) (MB-1 membrane glycoprotein) (Surface-IgM-
associated protein) (Membrane-bound immunoglobulin associated
protein) (CD79a).
DE CD79A OR IGA OR MB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92166394; PubMed=1538135;
 RA Ha H.J., Kubagawa H., Burrows P.D.;
 RT "Molecular cloning and expression pattern of a human gene homologous
 RL to the murine mb-1 gene.";
 RN J. Immunol. 148:1526-1531(1992).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92105765; PubMed=1723978;
 RA Yu L.M., Chang T.W.;
 RT "Human mb-1 gene: Complete cDNA sequence and its expression in B
 RL cells bearing membrane Ig of various isotypes.";
 RN J. Immunol. 148:633-637(1992).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93009083; PubMed=1395095;
 RA Leduc I., Preud'Homme J.L., Cogne M.;
 RT "Structure and expression of the mb-1 transcript in human lymphoid
 RL cells.";
 RN Clin. Exp. Immunol. 90:141-146(1992).
 [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347937; PubMed=1639443;
 RA Flawinkel H., Reth M.;
 RT "Molecular cloning of the Ig-alpha subunit of the human B-cell
 RL antigen receptor complex.";
 RN Immunogenetics 36:266-269(1992).
 [5]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Tonsil;
 RL MEDLINE=92288825; PubMed=1534761;
 RA Mueller B.S., Cooper L., Tertorst C.;
 RT "Cloning and sequencing of the cDNA encoding the human homologue of
 RL the murine immunoglobulin-associated protein B29.";
 RN Eur. J. Immunol. 22:1621-1625(1992).
 [6]
 RP SEQUENCE FROM N.A. (SEQUENCE SHORT).
 RA Koyama M., Nakamura T.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 [7]
 RP PRELIMINARY SEQUENCE OF 190-226 FROM N.A.
 RX MEDLINE=89091088; PubMed=2463161;
 RA Sakaguchi N., Kashiwamura S., Kinoto M., Thalmann P., Melchers F.;
 RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-
 RL like structural properties.";
 RN EMBO J. 7:3457-3464(1988).
 [8]
 RP SEQUENCE OF 33-52.
 RX MEDLINE=94239426; PubMed=7514267;
 RA Vasile S., Coligan J.E., Yoshida M., Seon B.K.;
 RT "Isolation and chemical characterization of the human B29 and mb-1
 RL proteins of the B cell antigen receptor complex.";
 RN Mol. Immunol. 31:419-427(1994).
 CC -1- FUNCTION: Associated to surface Igm-receptor; may be involved in
 CC signal transduction.
 CC -1- SUBUNIT: Heterodimer of alpha and beta chains; disulfide-linked.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Fl1912-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Fl1912-2; Sequence=VSP_002476;
 CC -1- TISSUE SPECIFICITY: B-cells.
 CC -1- PTM: Phosphorylated on tyrosine; upon B-cell activation.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- DATABASE: NAME=PROW; NOTE=CD Guide CD79a entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd79a.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U05259; AAA20495.1; -;
 DR EMBL; M86921; AAA59557.1; -;
 DR EMBL; X13451; CRA31802.1; ALT_SEQ.
 DR EMBL; S75217; AAB20812.1; -;
 DR EMBL; S46706; AAB23558.1; -;
 DR EMBL; M80462; AAA59556.1; -;
 DR EMBL; X83440; CAA58523.1; -;
 DR EMBL; M74721; AAA60270.1; -;
 DR PIR; I54539; A46477.
 DR Genew; HGNC:1698; CD79A.
 DR MIM; 112205; -;
 DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; TAS.
 DR GO; GO:0006952; P:defense response; TAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF00047; IG; 1.
 DR Pfam; PF02189; ITAM; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00077; ITAM; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 DR Immunoglobulin domain; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 226
 FT B-CELL ANTIGEN RECEPTOR COMPLEX
 FT ASSOCIATED PROTEIN ALPHA-CHAIN.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE.
 FT POTENTIAL.
 FT INTERCHAIN (WITH C-136 IN BETA CHAIN)
 FT (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT GTLIQNVKSHGIGYVCRVQEGNESYQQSCGYLVRVQ
 FT -> E (in isoform Short).
 FT /FtId=VSP_002476.
 FT V -> I (IN REF. 4).
 FT CONFLICT 69 69
 FT SEQUENCE 226 AA; 25038 MW; 6B5B837409969292 CRC64;
 Query Match 13.9%; Score 88; DB 1; Length 226;
 Best Local Similarity 29.9%; Pred. No. 0.061;
 Matches 38; Conservative 13; Mismatches 50; Indels 26; Gaps 7;
 QY 1 LMWSQ-PEIRTEGSSAFPCSFNASQGLAISGFWFRDEVVPGKEVRNGT--PEPRG 57
 DB 33 LMHEKVPASLAVSLGSDAHFQCPHNSNN----ANVTWVR--VLHG----NYTWPFPELG 82
 QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGNGTRLVVEKEHPQ-- 115
 DB 83 P-----GEPDNGTLLIQNVKSHGIGYVCRVQEGNESYQQSCGYLVRVQPPRPF 133
 QY 116 --LGAGT 120
 DB 134 LDMGEGT 140
 RESULT 9
 ID BUTY_MOUSE STANDARD; PRT; 524 AA.
 AC Q62556; P97392;
 DT 01-NOV-1997 (Rel. 35, Created)

RT "Complete sequence determination of the mouse and human CTLA4 gene
 RT loci: cross-species DNA sequence similarity beyond exon borders.";
 RL Genomics 60:341-355 (1999).
 RN [6]
 RN FUNCTION.
 RX MEDLINE=91341416; PubMed=1714933;
 RA Linsley P.S., Brady W., Urnes M., Griensmaire L.S., Danle N.K.,
 RA Ledbetter J.A.;
 RT "CTLA-4 is a second receptor for the B cell activation antigen B7";
 RL J. Exp. Med. 174:561-569 (1991).
 RN [7]
 RN STRUCTURE BY NMR OF 37-165.
 RX MEDLINE=97372889; PubMed=9228944;
 RA Metzler W.J., Bajjorath J., Fenderson W., Shaw S.Y., Constantine K.L.,
 RA Naemura J., Leytze G., Peach R.J., Lavoie T.B., Mueller L.,
 RA Linsley P.S.;
 RT "Solution structure of human CTLA-4 and delineation of a CD80/CD86
 RT binding site conserved in CD28";
 RL Nat. Struct. Biol. 4:527-531 (1997).
 RN [8]
 RN VARIANT ALA-17, AND ASSOCIATION WITH IDDM12.
 RX MEDLINE=97402209; PubMed=9259273;
 RA Marron M.P., Raffel L.J., Garchon H.-J., Jacob C.O., Serrano-Rios M.,
 RA Marton M.P., Raffel L.J., Garchon H.-J., Zhang Z.-X.,
 RA Goldstein D.R., Tao Y.-W., Beaurain G., Bach J.-F., Huang H.-S.,
 RA Luo D.-F., Zeidler A., Rottier J.I., Yang M.C.K., Modilevsky T.,
 RA Maclaren N.K., She J.-X.;
 RT "Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4
 RT polymorphisms in multiple ethnic groups";
 RL Hum. Mol. Genet. 6:1275-1282 (1997).
 RN [9]
 RN POLYMORPHISM, AND ASSOCIATION WITH COELIAC DISEASE.
 RX MEDLINE=99205840; PubMed=10189842;
 RA Djilali-Saiah I., Schmitz J., Harfouch-Hamoud E., Mougnot J.-F.,
 RA Bach J.-F., Caillaud-Zucman S.;
 RT "CTLA-4 gene polymorphism is associated with predisposition to coeliac
 RT disease";
 RL Gut 43:187-189 (1998).
 RN [10]
 RN VARIANT ALA-17, AND ASSOCIATION WITH TAO.
 RX MEDLINE=99402177; PubMed=10475192;
 RA Vaidya B., Imrie H., Perros P., Dickinson J., McCarthy M.I.,
 RA Kendall-Taylor P., Pearce S.H.S.;
 RT "Cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphism confers
 RT susceptibility to thyroid associated orbitopathy";
 RL Lancet 334:743-744 (1999).
 RN [11]
 RN VARIANT ALA-17, AND ASSOCIATION WITH GRD.
 RX MEDLINE=20385252; PubMed=10924276;
 RA Chistyakov D.A., Savost'yanov K.V., Turakulov R.I., Petunina N.A.,
 RA Trukhina L.V., Kudanova A.V., Balabolkin M.I., Nosikov V.V.;
 RT "Complex association analysis of Graves disease using a set of
 RT polymorphic markers";
 RL Mol. Genet. Metab. 70:214-218 (2000).
 RN [12]
 RN VARIANT ALA-17.
 RX MEDLINE=20395844; PubMed=10903931;
 RA Deng Z., Morse J.H., Slager S.L., Cuervo N., Moore K.J., Venetos G.,
 RA Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,
 RA Knowles J.A.;
 RT "Familial primary pulmonary hypertension (gene PPH1) is caused by
 RT mutations in the bone morphogenetic protein receptor-II gene";
 RL Am. J. Hum. Genet. 67:737-744 (2000).
 CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
 CC (CD80) AND B7-2 (CD86).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC lymphoid tissues.
 CC -1- POLYMORPHISM: The variant Ala-17 is associated with an increased
 CC risk for autoimmune disorders as Graves disease (GRD), type I
 CC insulin-dependent diabetes mellitus (IDDM12), and thyroid-
 CC associated orbitopathy (TAO). The variant Thr-17 is associated
 CC with predisposition to coeliac disease, a gluten sensitive

enteropathy characterized by small bowel mucosal atrophy.
 -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 -1- DATABASE: NAME=PROW; NOTE=CD guide CD152 entry;
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd152.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF411058; AAL40932.1; -
 CC EMBL; AF41120; AAL07473.1; -
 CC EMBL; M74363; AAA52127.1; -
 CC EMBL; M37245; AAA52773.1; -
 CC EMBL; M37243; AAA52773.1; JOINED.
 CC EMBL; M37244; AAA52773.1; JOINED.
 CC EMBL; AF42144; AAF02499.1; -
 CC PIR; S08614; S08614.
 CC PDB; 1AHL; 15-APR-98.
 CC PDB; 1H6E; 28-NOV-01.
 CC PDB; 1I85; 04-APR-01.
 CC PDB; 1I8L; 04-APR-01.
 CC Genew; HGNC:2505; CTLA4.
 CC MIM; 123890; -
 CC GO; GO:0005887; C: integral to plasma membrane; TAS.
 CC GO; GO:0006955; P: immune response; TAS.
 CC InterPro; IPR008096; CTLA4.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003596; IG V.
 CC PRINTS; PR01720; CTLANTIGEN4.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00406; IGV; 1.
 CC PROSITE; PS00835; IG LIKE; FALSE NEG.
 CC Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal;
 KW 3D-structure; Polymorphism.
 FT SIGNAL 1 35
 FT CHAIN 36 223
 FT DOMAIN 36 161
 FT TRANSMEM 162 187
 FT DOMAIN 188 223
 FT DOMAIN 39 140
 FT DISULFID 58 129
 FT DISULFID 85 103
 FT CARBOHYD 113 113
 FT VARIANT 17 17
 FT CONFLICT 147 147
 FT STRAND 45 47
 FT TURN 50 51
 FT STRAND 53 59
 FT STRAND 70 76
 FT STRAND 81 88
 FT TURN 91 92
 FT TURN 99 100
 FT STRAND 104 108
 FT TURN 109 110
 FT STRAND 111 117
 FT TURN 121 123
 FT STRAND 125 132
 FT TURN 137 138
 FT STRAND 142 143
 FT STRAND 147 150
 FT SEQUENCE 223 AA; 24656 MW; 6F9466FB2E139A5A CRC64;

Query Match 13.8%; Score 87.5; DB 1; Length 223;

Best Local Similarity 27.4%; Fred. No. 0.068; 56; Indels 19; Gaps 4;

Matches 34; Conservative 15; Mismatches 15; Indels 19; Gaps 4;

QY 3 VSQPEIRTEGSSAFPCSFNAGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLAPL 62

```
Db 40 VAQPAVLASSRGASVCEY-ASPGKATEVRVTVLRQADSQVTECAATYMGNELT-- 96
Qy 63 ASSRFLHD-----HQELHIRDVRGDASIYCRVEVL---GLGVGTNGRLVVE 110
Db 97 ----FLDSDICTGTSGNQVNLTIQGLRAMDTGLYICKVELMYPFPYILGNGTQIYVI 152
Qy 111 KEHP 114
Db 153 DPEP 156

RESULT 11
MOG_BOVIN
ID_MOG_BOVIN STANDARD; PRT; 246 AA.
AC P55803;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-70.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,
RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex."
RN Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
[2]
RP SEQUENCE OF 29-36.
RC TISSUE=Brain;
RX MEDLINE=93382604; PubMed=8371836;
RA Birling M.C., Rousset G., Nussbaum J.-L.;
RT "Biochemical and immunohistochemical studies with specific polyclonal
RT antibodies directed against bovine myelin/oligodendrocyte
RT glycoprotein."
RN Neurochem. Res. 18:937-945 (1993).
CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -!- SUBUNIT: May form homodimers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L21757; -; NOT_ANNOTATED_CDS.
DR PIR: A4712; A4712.
DR HSRF: Q13740; 1KZC.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003596; Ig_v.
```

```
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00835; IG LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 246
FT DOMAIN 29 153
FT TRANSMEM 154 174
FT DOMAIN 175 209
FT TRANSMEM 210 230
FT DOMAIN 231 246
FT DISULFID 52 126
FT CARBOHYD 59 59
SQ SEQUENCE 246 AA; 363C76AB1A33DE41 CRC64;

Query Match 13.8%; Score 87.5; DB 1; Length 246;
Best Local Similarity 34.0%; Pred. No. 0.076;
Matches 32; Conservative 9; Mismatches 34; Indels 19; Gaps 6;

Qy 9 IRTLEGSAFLPCSFNASQQLAIG-SVTWFRDEWVPGKEV-----RNG-----TPEFR 56
Db 40 IRALVGVDELPC--RISPGKNATGMEVGNVYP---PFSRVVHLYRNGKQDDEQAPEYR 94
Qy 57 GRLAPLASSRFLHDHQELHIRDVRGDASIYVC 90
Db 95 GRTQLKEI--IGEKVTLIRNVRFSDEGGFTC 126

RESULT 12
MOG_HUMAN
ID_MOG_HUMAN STANDARD; PRT; 247 AA.
AC Q16553; Q00713; Q00714; Q00715; Q13054; Q13055; Q14855; Q92891;
AC Q92892; Q92893; Q92894; Q92895; Q93053; Q96KU9; Q96KU0; Q96605;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=93310943; PubMed=7790876;
RA Hilton A.A., Slavin A.J., Hilton D.J., Bernard C.C.A.;
RT "Characterization of cDNA and genomic clones encoding human myelin
RT oligodendrocyte glycoprotein."
RL J. Neurochem. 65:309-318 (1995).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95054056; PubMed=7964757;
RA Pham-Dinh D., Allingant B., Ruberg M., della Gaspera B.,
RA Nussbaum J.-L., Dautigny A.;
RT "Characterization and expression of the cDNA coding for the human
RT myelin/oligodendrocyte glycoprotein."
RL J. Neurochem. 63:2353-2356 (1994).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96015053; PubMed=8530032;
RA Roth M.-P., Malfroy L., Offer C., Sevin J., Enault G., Borot N.,
RA Pontarotti P., Coppin H.;
RT "The human myelin oligodendrocyte glycoprotein (MOG) gene: complete
RT nucleotide sequence and structural characterization."
RL Genomics 28:241-250 (1995).
[4]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96115584; PubMed=8666381;
RA Pham-Dinh D., della Gaspera B., de Rosbo N.K., Dautigny A.;
RT "Structure of the human myelin/oligodendrocyte glycoprotein gene and
RT multiple alternative spliced isoforms."
RT
```

Genomics 29:345-352 (1995).

[5] SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RA Ballenthin P.A., Gardiner M.V.; to the EMBL/GenBank/DBSJ databases.

[6] SEQUENCE FROM N.A. (ISOFORMS 1; 5 AND 7).

RA Griffiths C.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.

[7] STRUCTURE BY NMR OF 54-84.

RP MEDLINE=97354172; PubMed=9210466;

RA Albouze-Abou S., Wilson J.C., Bernard C.C.A., von Itzstein M.;

RT "A conformational study of the human and rat encephalitogenic myelin

RT oligodendrocyte glycoprotein peptides 35-55.";

RL Eur. J. Biochem. 246:59-70 (1997).

CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in

CC completion and/or maintenance of the myelin sheath and in cell-

CC cell communication.

CC -!- SUBUNIT: May form homo- or heterodimers between the different

CC isoforms.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (isoforms 1 and

CC 5); type I membrane protein (other isoforms) (Potential).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=9;

CC Comment=Additional isoforms seem to exist;

CC Name=1; Synonyms=Alpha-1;

CC IsoId=Q16653-1; Sequence=Displayed;

CC Name=2; Synonyms=Alpha-2;

CC IsoId=Q16653-2; Sequence=VSP_002543;

CC Name=3; Synonyms=Alpha-3;

CC IsoId=Q16653-3; Sequence=VSP_002542;

CC Name=4; Synonyms=Alpha-4;

CC IsoId=Q16653-4; Sequence=VSP_002539;

CC Name=5; Synonyms=Beta-1;

CC IsoId=Q16653-5; Sequence=VSP_002545;

CC Name=6; Synonyms=Beta-2;

CC IsoId=Q16653-6; Sequence=VSP_002543; VSP_002545;

CC Name=7; Synonyms=Beta-3;

CC IsoId=Q16653-7; Sequence=VSP_002542; VSP_002545;

CC Name=8; Synonyms=Beta-4;

CC IsoId=Q16653-8; Sequence=VSP_002544; VSP_002545;

CC Name=9;

CC IsoId=Q16653-9; Sequence=VSP_002540; VSP_002541;

CC Note=Not functionally active. May be expressed at low level in

CC the adult;

CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is

CC localized on the surface of myelin and oligodendrocyte cytoplasmic

CC membranes.

CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG

CC family.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)

CC with oligodendrocyte-myelin glycoprotein (OMG).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X74511; CAAS2617.1; -

DR EMBL; Z48051; CAAB8109.1; -

DR EMBL; U18840; AAC50361.1; -

DR EMBL; U18843; AAC50362.1; -

DR EMBL; U18798; AAC50876.1; -

DR EMBL; U18799; AAC50877.1; -

DR EMBL; U18800; AAC36870.1; -

DR EMBL; U18801; AAC50878.1; -

DR EMBL; U18803; AAC50879.1; -

DR EMBL; U64564; AAB08088.1; -

DR EMBL; U64565; AAB08089.1; -

DR EMBL; U64566; AAB08090.1; -

DR EMBL; U64567; AAB08091.1; -

DR EMBL; U64568; AAB08092.1; -

DR EMBL; U64569; AAB08093.1; -

DR EMBL; U64570; AAB08094.1; -

DR EMBL; U64571; AAB08095.1; ALT_SEQ.

DR EMBL; AL050328; CAB89267.1; -

DR EMBL; AL050328; CAB89269.1; -

DR EMBL; AL050328; CAB89270.1; -

DR EMBL; AL050328; CAB89271.1; -

DR Genew; HGNC:7197; MOG.

DR MIN; 159465; -

DR GO; GO:0007417; P:central nervous system development; TAS.

DR InterPro; IPR007110; Ig-like.

DR Pfam; PF00047; ig; 1.

DR PROSITE; PS00835; IG_LIKE; 1.

DR Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal;

KW Alternative splicing.

KW SIGNAL 1 29

FT CHAIN 30 247

FT DOMAIN 30 154

FT TRANSMEM 155 175

FT DOMAIN 176 210

FT TRANSMEM 211 231

FT DOMAIN 232 247

FT DOMAIN 32 145

FT DISULFID 53 127

FT CARBOHYD 60 60

FT VARSPLIC 30 145

FT VARSPLIC 198 203

FT VARSPLIC 204 247

FT VARSPLIC 198 236

FT VARSPLIC 198 236

FT VARSPLIC 198 236

FT VARSPLIC 198 243

FT VARSPLIC 244 247

FT CONFLICT 171 171

FT SEQUENCE 247 AA; 28179 MW; 847601FE5937AB0C CRC64;

Query Match 13.8%; Score 87.5; DB 1; Length 247;

Best Local Similarity 34.0%; Pred. No. 0.076;

Matches 32; Conservative 9; Mismatches 34; Indels 19; Gaps 6;

QY 9 IRTLEGSSAFPCSFNASQGLAIG-SVTWFRDEVVPGKEV-----RNG-----TPEPR 56

DB 41 IRLVGVDEVLPCC--RISPGKATGMEVGRIP---PFSRVVHLYRNGKQDGDQAPERY 95

QY 57 GRPLAPASSRFLHDHQAELHVRDVRGHDAIYVC 90

DB 96 GRTELLAKDA--IGEGKVTLRIRNVRFSDGGFTC 127

RESULT 13

CTLA4_PIG STANDARD; PRT; 223 AA.

ID CTL4_PIG

AC Q9NMV7;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-

DE associated antigen 4) (CTLA-4).

GN CTLA4.

OS Sus scrofa (Pig).

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 16.1273 Seconds
(without alignments)
613.452 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 MAWMLLLILIMVHPSGALW.....GTHCHSDGPRGVIPEPCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	11.0	223	1	CTL4_HUMAN
2	111.5	10.9	226	1	C79A_HUMAN
3	108.5	10.6	223	1	CTL4_MOUSE
4	106	10.4	132	1	TVB2_MOUSE
5	103.5	10.1	223	1	CTL4_PIG
6	103.5	10.1	223	1	CTL4_RABIT
7	100.5	9.9	1694	1	SN_MOUSE
8	99	9.7	467	1	SIL5_MOUSE
9	98.5	9.7	526	1	BUTY_HUMAN
10	98	9.6	245	1	MOG_RAT
11	96.5	9.5	246	1	MOG_MOUSE
12	96	9.4	131	1	LV6E_HUMAN
13	95	9.3	524	1	BUTY_MOUSE
14	93.5	9.2	247	1	MOG_HUMAN
15	93	9.1	246	1	MPPO_HETFR
16	92.5	9.1	403	1	CD33_MOUSE
17	92	9.0	215	1	CTB2_HUMAN
18	90.5	8.9	246	1	MOG_BOVIN
19	89	8.7	142	1	VPB2_MOUSE
20	89	8.7	526	1	BUTY_BOVIN
21	89	8.7	628	1	LU_HUMAN
22	89	8.7	1709	1	SN_HUMAN
23	88.5	8.7	1197	1	CAM1_BRARE
24	88	8.6	463	1	SIL9_HUMAN
25	88	8.6	579	1	ASO_CUCMA
26	85.5	8.4	3707	1	PCBM_MOUSE
27	85	8.3	2132	1	PCCA_MOUSE
28	84	8.2	133	1	TVB2_HUMAN
29	84	8.2	2415	1	PCCA_HUMAN
30	83	8.1	134	1	TVB7_MOUSE
31	83	8.1	270	1	BASI_RABIT
32	83	8.1	587	1	ASO_CUCA
33	83	8.1	2333	1	PCCA_CANFA

ALIGNMENTS

RESULT 1

CTL4_HUMAN					
ID	CTL4_HUMAN	STANDARD	PRT	223	AA
AC	P16410; Q96F43; Q9UKN9;				
DT	01-AUG-1990 (Rel. 15, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-associated antigen 4) (CTLA-4) (CD152 antigen).				
DE	CTLA4 OR CD152.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21601151; PubMed=11735222;				
RA	Ling V., Wu P.W., Finnerty H.F., Agostino M.J., Graham J.R., Chen S., Jussiff J., Fisk G.J., Miller C.P., Collins M.;				
RA	"Assembly and annotation of human chromosome 2q33 sequence containing the CD28, CTLA4, and ICOS gene cluster: analysis by computational, comparative, and microarray approaches.";				
RL	Genomics 78:155-168(2001).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Wu P.W., Ling V.;				
RA	"Full length sequence of hCTLA4 cDNA.";				
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.				
RN	[3]				
RP	SEQUENCE OF 1-37 FROM N.A.				
RX	MEDLINE=91318145; PubMed=1713603;				
RA	Harper K., Balzano C., Novier E., Mattei M.-G., Luciani M.F., Golstein P.;				
RA	"CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mouse and human as to sequence, message expression, gene structure, and chromosomal location.";				
RL	J. Immunol. 147:1037-1044(1991).				
RN	[4]				
RP	SEQUENCE OF 38-223 FROM N.A.				
RP	TISSUE=Lymphocytes;				
RC	MEDLINE=89120925; PubMed=3220103;				
RA	Dariavach P., Mattei M.-G., Golstein P., Lefranc M.-P.;				
RA	"Human Ig superfamily CTLA-4 gene: chromosomal localization and identity of protein sequence between murine and human CTLA-4 cytoplasmic domains.";				
RL	Eur. J. Immunol. 18:1901-1905(1988).				
RN	[5]				
RP	SEQUENCE OF 140-223 FROM N.A., AND TISSUE SPECIFICITY.				
RP	MEDLINE=99425274; PubMed=10493833;				
RA	Ling V., Wu P.W., Finnerty H.F., Sharpe A.H., Gray G.S., Collins M.;				
RA	"Complete sequence determination of the mouse and human CTLA4 gene loci: cross-species DNA sequence similarity beyond exon borders.";				
RL	Genomics 60:341-355(1999).				
RN	[6]				
RP	FUNCTION.				
RX	MEDLINE=91341416; PubMed=1714933;				

Q951h0 pan troglod
P43121 homo sapien
P78310 homo sapien
O15389 homo sapien
P20916 homo sapien
P07897 rattus norv
P40293 bos taurus
P37084 cucurbita p
P97792 mus musculu
P13372 mus musculu
P06316 homo sapien
Q9nyz4 homo sapien

RA Linsley P.S., Brady W., Urnes M., Griosmaire L.S., Damle N.K.,
 RA Ledbetter J.A.;
 RA "CTLA-4 is a second receptor for the B cell activation antigen B7";
 RL J. Exp. Med. 174:561-569(1991).
 [7]
 RN STRUCTURE BY NMR OF 37-165.
 RX MEDLINE=97372899; PubMed=9228944;
 RA Metzler W.J., Bajjorath J., Fenderson W., Shaw S.Y., Constantine K.L.,
 RA Nemura J., Leytze G., Peach R.J., Lavoie T.B., Mueller L.,
 RA Linsley P.S.;
 RA "Solution structure of human CTLA-4 and delineation of a CD86/CD86
 RT binding site conserved in CD28";
 RL Nat. Struct. Biol. 4:527-531(1997).
 [8]
 RN VARIANT ALA-17, AND ASSOCIATION WITH IDDM12.
 RX MEDLINE=97402209; PubMed=9259273;
 RA Marron M.P., Raffel L.J., Garchon H.-J., Jacob C.O., Serrano-Rios M.,
 RA Martinez Larrad M.F., Teng W.-P., Park Y., Zhang Z.-X.,
 RA Goldstein D.R., Tao Y.-W., Beaurain G., Bach J.-F., Huang H.-S.,
 RA Luo D.-F., Zeidler A., Rother J.I., Yang M.C.K., Modilevsky T.,
 RA MacLaren N.K., She J.-X.;
 RA "Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4
 RT polymorphisms in multiple ethnic groups";
 RL Hum. Mol. Genet. 6:1275-1282(1997).
 [9]
 RN POLYMORPHISM, AND ASSOCIATION WITH COELIAC DISEASE.
 RX MEDLINE=99205840; PubMed=10189842;
 RA Djilali-Saiah I., Schmitz J., Harfouch-Hamoud E., Mougnot J.-F.,
 RA Bach J.-F., Caillaud-Zucman S.;
 RA "CTLA-4 gene polymorphism is associated with predisposition to coeliac
 RT disease";
 RL Lancet 354:743-744(1999).
 [10]
 RN VARIANT ALA-17, AND ASSOCIATION WITH TAO.
 RX MEDLINE=99402177; PubMed=10475192;
 RA Valdiva B., Imrie H., Perros P., Dickinson J., McCarthy M.I.,
 RA Kendall-Taylor P., Pearce S.H.S.;
 RA "Cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphism confers
 RT susceptibility to thyroid associated orbitopathy";
 RL Lancet 354:743-744(1999).
 [11]
 RN VARIANT ALA-17, AND ASSOCIATION WITH GRD.
 RX MEDLINE=20385252; PubMed=10924276;
 RA Chistakov D.A., Savost'yanov K.V., Turakulov R.I., Petunina N.A.,
 RA Trukhina L.V., Kudanova A.V., Balabolkin M.I., Nosikov V.V.;
 RA "Complex association analysis of Graves disease using a set of
 RT polymorphic markers";
 RL Mol. Genet. Metab. 70:214-218(2000).
 [12]
 RN VARIANT ALA-17.
 RX MEDLINE=20395844; PubMed=10903931;
 RA Deng Z., Morse J.H., Slager S.L., Quervo N., Moore K.J., Venetos G.,
 RA Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,
 RA Knowles J.A.;
 RA "Familial primary pulmonary hypertension (gene PPH1) is caused by
 RT mutations in the bone morphogenetic protein receptor-II gene";
 RL Am. J. Hum. Genet. 67:737-744(2000).
 CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
 CC (CD80) AND B7-2 (CD86).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC lymphoid tissues.
 CC -1- POLYMORPHISM: The variant Ala-17 is associated with an increased
 CC risk for autoimmune disorders as Graves disease (GRD), type I
 CC insulin-dependent diabetes mellitus (IDDM12), and thyroid-
 CC associated orbitopathy (TAO). The variant Thr-17 is associated
 CC with predisposition to coeliac disease, a gluten sensitive
 CC enteropathy characterized by small bowel mucosal atrophy.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD152 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd152.htm";
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF411058; AAL40932.1; -;
 DR EMBL; AF414120; AAL07473.1; -;
 DR EMBL; M74363; AAA52127.1; -;
 DR EMBL; M37245; AAA52773.1; -;
 DR EMBL; M37243; AAA52773.1; JOINED.
 DR EMBL; M37244; AAA52773.1; JOINED.
 DR EMBL; AF142144; AAF02499.1; -;
 DR PIR; S08614; S08614.
 DR PDB; 1AHL1; 15-APR-98.
 DR PDB; 1H6E; 28-NOV-01.
 DR PDB; 1I85; 04-APR-01.
 DR PDB; 1I8L; 04-APR-01.
 DR Genew; HGNC:2505; CTLA4.
 DR MIM; 123890; -;
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0006955; P: immune response; TAS.
 DR InterPro; IPR008096; CTLA4.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR PRINTS; PR01720; CTLANTIGEN4.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00406; IGv; 1.
 DR PROSITE; PS50835; IG_LIKE; FALSE NEG.
 DR Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal;
 KW 3D-structure; Polymorphism.
 FT SIGNAL 1 35
 FT CHAIN 36 223
 FT DOMAIN 36 161
 FT TRANSMEM 162 187
 FT DOMAIN 188 223
 FT DOMAIN 39 140
 FT DISULFID 58 129
 FT DISULFID 85 103
 FT CARBOHYD 113 113
 FT VARIANT 17 17
 FT CONFLICT 147 147
 FT STRAND 45 47
 FT TURN 50 51
 FT STRAND 53 59
 FT STRAND 70 76
 FT STRAND 81 88
 FT TURN 91 92
 FT TURN 99 100
 FT STRAND 104 108
 FT TURN 109 110
 FT STRAND 111 117
 FT TURN 121 123
 FT STRAND 125 132
 FT TURN 137 138
 FT STRAND 142 143
 FT STRAND 147 150
 FT SEQUENCE 223 AA; 24656 MW; 6F9466FB2E139A5A CRC64;
 SQ
 Query Match 11.0%; Score 112; DB 1; Length 223;
 Best Local Similarity 28.0%; Pred. No. 0.0038;
 Matches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;
 QY 6 LLLIMWPGSC-ALWYSOPPEIRLEGSASFPCSNAGSLAIGSVTFWFRDEVVPGK 64
 DB 24 LLFLFLFPVFCWHAQVPAVILASSRGIAFVCEY-ASPKATEVTVVLRDQSVT 82
 QY 65 EYRNGTPEFRGRGLAPLASSRFLHD-----HQAELHIRDVRGHDASIYCRVYL-- 113
 DB 83 EYCAATYMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVLMYP 136

QY 114 -GLGVCTGNGTRLVVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156
 Db 137 PPYLIGNGTQIYVIDPEPCDPSDFLLWLAASGLFFYSFLTAV 184

RESULT 2
 C79A HUMAN STANDARD; PRT; 226 AA.
 ID C79A HUMAN STANDARD; PRT; 226 AA.
 AC P11912;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE B-cell antigen receptor complex associated protein alpha-chain
 DE precursor (Ig-alpha) (MB-1 membrane glycoprotein) (Surface-IGM-
 DE associated protein) (Membrane-bound immunoglobulin associated
 DE protein) (CD79a).
 GN CD79A OR IGA OR MBL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9216394; PubMed=1538135;
 RA Ha H.J.; Kubagawa H.; Burrows P.D.;
 RT "Molecular cloning and expression pattern of a human gene homologous
 RT to the murine mb-1 gene.";
 RL J. Immunol. 148:1526-1531(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92105765; PubMed=1729378;
 RA Yu L.M.; Chang T.W.;
 RT "Human mb-1 gene: complete cDNA sequence and its expression in B
 RT cells bearing membrane Ig of various isotypes.";
 RL J. Immunol. 148:633-637(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93009083; PubMed=1395095;
 RA Leduc I.; Preud'Homme J.L.; Cogne M.;
 RT "Structure and expression of the mb-1 transcript in human lymphoid
 RT cells.";
 RL Clin. Exp. Immunol. 90:141-146(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92347937; PubMed=1639443;
 RA Flawinkel H.; Reth M.;
 RT "Molecular cloning of the Ig-alpha subunit of the human B-cell
 RT antigen receptor complex.";
 RL Immunogenetics 36:266-269(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92289825; PubMed=1534761;
 RA Mueller B.S.; Cooper L.; Terhorst C.;
 RT "Cloning and sequencing of the cDNA encoding the human homologue of
 RT the murine immunoglobulin-associated protein B29.";
 RL Eur. J. Immunol. 22:1621-1625(1992).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RA Koyama M.; Nakamura T.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP PRELIMINARY SEQUENCE OF 190-226 FROM N.A.
 RX MEDLINE=89091088; PubMed=2463161;
 RA Sakaguchi N.; Kashiwamura S.; Kimoto M.; Thalmann P.; Melchers F.;
 RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-
 RT like structural properties.";
 RL EMBO J. 7:3457-3464(1988).
 RN [8]
 RP SEQUENCE OF 33-52.
 RX MEDLINE=94239426; PubMed=7514267;
 RA Vasile S.; Coligan J.E.; Yoshida M.; Seon B.K.;

"Isolation and chemical characterization of the human B29 and mb-1
 proteins of the B cell antigen receptor complex.";
 Mol. Immunol. 31:419-427(1994).
 CC -|- FUNCTION: Associated to surface IGM-receptor; may be involved in
 CC signal transduction.
 CC -|- SUBUNIT: Heterodimer of alpha and beta chains; disulfide-linked.
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P11912-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P11912-2; Sequence=VSP_002476;
 CC -|- TISSUE SPECIFICITY: B-cells.
 CC -|- PTM: Phosphorylated on tyrosine; upon B-cell activation.
 CC -|- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -|- DATABASE: NAME=PROW; NOTE=CD guide CD79a entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd79a.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U05259; AAA20495.1; -;
 CC EMBL; M86921; AAA59557.1; -;
 CC EMBL; X13451; CAA31802.1; ALT_SEQ.
 CC EMBL; S75217; AAB20812.1; -;
 CC EMBL; M80462; AAB23558.1; -;
 CC EMBL; M80462; AAB23558.1; -;
 CC EMBL; X83540; CAA58523.1; -;
 CC EMBL; M74721; AAA60270.1; -;
 CC PIR; I54539; A46477.
 CC Gene; HGNC:1698; CD79A.
 CC MIM; 112205; -;
 CC GO; GO:0007166; P:cell surface receptor linked signal transdu. .; TAS.
 CC GO; GO:0006952; P:defense response; TAS.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003110; ITAM.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF02189; ITAM; 1.
 CC SMART; SM00077; ITAM; 1.
 CC PROSITE; PS50635; IG_LIKE; 1.
 CC Immunoglobulin domain; Transmembrane; Glycoprotein; Signal;
 KW Phosphorylation; Alternative splicing.
 FT SIGNAL 1 32
 FT CHAIN 33 226
 FT B-CELL ANTIGEN RECEPTOR COMPLEX
 FT ASSOCIATED PROTEIN ALPHA-CHAIN.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE C2-TYPE.
 FT POTENTIAL.
 FT INTERCHAIN (WITH C-136 IN BETA CHAIN)
 FT (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT GTLIQNVKSHGIVYVCRVQEGNESYQQSCGYLRVQ
 FT -> E (in isoform Short).
 FT /FTid=VSP_002476.
 FT V -> I (IN REF. 4).
 FT CONFLICT 69 69
 FT SEQUENCE 226 AA; 25038 MW; 6BSB83740969292 CRC64;
 SQ

Query Match 10.9%; Score 111.5; DB 1; Length 226;

Best Local Similarity 28.0%; Pred. No. 0.0043;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILIMVHPSGSCALWVSQ-PEIRITLGGSAFLPCSFNASQRLAIGSVTWFRDEVVP 62
DB 18 LFLSLAYLGGCOALMHWKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWVR--VLH 71
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDHQAELHIRDVGRHDASIVYCVREVLGLGVGTG 120
DB 72 G-----NYTWPEFLGP-----GDPNGTILIQNVKSHGIVYCVRQENESYQOS 118
QY 121 NGRPLVVEKHPQ-----LGATV-LLRAGFYAVSFLSVAVGTVYVQ 163
DB 119 CGTYLRVRQPPRPFDFMGEGTKNRITAEIGIILFCVAVPGTLLIFR 166

RESULT 3
CTLA4 MOUSE
ID CTLA4 MOUSE STANDARD; PRT; 223 AA.
AC P09793; Q9QZ27;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-associated antigen 4) (CTLA-4).
GN CTLA4 OR CD152.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97258259; PubMed=3496540;
RA Brunet J.-F., Denizot F., Luciano M.-F., Roux-Dosseto M., Suzan M., Mattei M.-G., Golstein P.;
RT "A new member of the immunoglobulin superfamily -- CTLA-4.";
RL Nature 328:267-270(1987).
RN [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX STRAIN=129/SvJ;
RX MEDLINE=99425274; PubMed=10493833;
RT Ling V., Wu P.W., Finnerty H.F., Sharpe A.H., Gray G.S., Collins M.;
RT "Complete sequence determination of the mouse and human CTLA4 gene loci: cross-species DNA sequence similarity beyond exon borders.";
RL Genomics 60:341-355(1999).
RN [3]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=91318145; PubMed=1713603;
RA Harper K., Balzano C., Rouvier E., Mattei M.-G., Luciano M.F., Golstein P.;
RT "CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mouse and human as to sequence, message expression, gene structure, and chromosomal location.";
RL J. Immunol. 147:1037-1044(1991).
CC -!- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1 (CD80) AND B7-2 (CD86).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest levels in lymphoid tissues.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X05719; CAA29191.1; --
DB EMBL; AF142145; AAP01489.1; --
DB EMBL; M74362; AAA37489.1; --
PIR; A29063; A29063.

DR HSP; P16410; 1RH1.
DR MGD; MGI:88556; Ctl4.
DR InterPro; IPR008096; Ctl4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR01720; CTLANTIGEN4.
DR SMART; SM00409; Ig; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 35
FT CHAIN 36 223
FT DOMAIN 36 161
FT TRANSMEM 162 187
FT DOMAIN 188 223
FT DOMAIN 36 145
FT DISULFID 58 129
FT DISULFID 85 103
FT CARBOHYD 108 108
FT CARBOHYD 113 113
FT CARBOHYD 145 145
FT CONFLICT 182 182
SQ SEQUENCE 223 AA; 24993 MW; 5318FAAF416F4685 CRC64;
Query Match 10.6%; Score 108.5; DB 1; Length 223;
Best Local Similarity 26.9%; Pred. No. 0.0079;
Matches 45; Conservative 26; Mismatches 79; Indels 17; Gaps 6;

QY 3 MLLILIMVHPSGSCALWVSQ-PEIRITLGGSAFLPCSFNASQRLAIGSVTWFRD--- 58
DB 22 FVALLTLFIPIVFEAIQVTPSVVLAASHGVAFPCPEYSPSHNTDEV-RVTVLRQTNDQ 80
QY 59 --EVVPGKEVNGTPEFRGLAPLASSRFLHDHQAELHIRDVGRHDASIVYCVREVL--- 113
DB 81 MTEVCATFTTEKNTVGELD--YPCSGTTF-NESRVNLTIOGLRAVDGLYLCKVELMYP 137
QY 114 GLGVGTGNGTRLVVEKEHPQLGAGTVL-----LLRAGFYAVSFLSVAV 156
DB 138 PFVGMGNGTQIYVIDPEPCDSDFLWILVAVSLGLFFYSFLVSAV 184

RESULT 4
TV2A2 MOUSE
ID TV2A2 MOUSE STANDARD; PRT; 132 AA.
AC P01739;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell receptor alpha chain V region 2B4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (CLONE TT11).
RX MEDLINE=85036634; PubMed=6548551;
RA Chien Y., Becker D.M., Lindsten T., Okamura M., Cohen D.I., Davis M.M.;
RT "A third type of murine T-cell receptor gene.";
RL Nature 312:31-35(1984).
PIR; A02015; RWSAV.
DR HSP; P01607; 1REI.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig.V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW T-cell; Receptor; Immunoglobulin domain; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 132
FT DOMAIN 21 113
FT DOMAIN 114 117
FT DOMAIN 118 132
FT T-CELL RECEPTOR ALPHA CHAIN V REGION 2B4.
FT V SEGMENT.
FT D SEGMENT.
FT J SEGMENT.

```

FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .).
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 14668 MW; CABBF6CF1DD3448B CRC64;

Query Match 10.4%; Score 106; DB 1; Length 132;
Best Local Similarity 25.8%; Pred. No. 0.0077;
Matches 40; Conservative 22; Mismatches 45; Indels 48; Gaps 9;

QY 3 WMLLLILMVHFGSCALWVS-----QPPEIRTLLEGSAFLPCSFNASQGRLAIGSVTW 55
Dy 12 WLL-----NWNSQNVQSPESLIVPEGARTSLNCTFSDSASQY-----FWW 55

QY 56 ERDEVVPGKEVR-----NGTPEFGRPLAPLASSRFLDHOAEHLHVRDVRGHDSIYVC 108
Dy 56 YRQH--SGKAPKALMIFNGEKE--EGRTIHLNKASLH-----FSLHSDQSDSALYLC 109

QY 109 RVEVLGLVGTGNGTRLVVKEHPOLGAGTVLLLR 143
Dy 110 AVTLYG-----GSGNKL-----FGTGTLLSVK 132

RESULT 5
CTL4_PIG STANDARD; PRT; 223 AA.
ID CTL4_PIG AC Q9MYX7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-
DE associated antigen 4) (CTLA-4).
GN CTLA4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Tachedjian M., Chaplin P.J., Scheerlinck J.-P.Y., Tennent J.M.;
RT "Molecular characterization and phylogenetic analysis of porcine
RT cytotoxic T-lymphocyte-associated antigen 4 (CTLA4).";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
CC (CD80) AND B7-2 (CD86) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF281633; AAP86981.1; -.
CC HSP; P16410; IAH1.
CC InterPro; IPR008096; CTLA4.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC PRINTS; PR01720; CTLANTIGEN4.
CC SMART; SM00409; IG; 1.
CC DR PROSITE; PS50835; IG LIKE; FALSE NEG.
CC Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 35
FT CHAIN 36 223
FT DOMAIN 162 187
FT TRANSMEM 188 223
FT CYTOTOXIC T-LYMPHOCYTE PROTEIN 4.
FT POTENTIAL.
FT DOMAIN 188 223
FT CYTOPLASMIC (POTENTIAL).
FT BY SIMILARITY.
FT DISULFID 58 129
FT DISULFID 85 103
FT CARBOHYD 113 113
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145

SQ SEQUENCE 223 AA; 24430 MW; 2DDB865D871D8921 CRC64;

Query Match 10.1%; Score 103.5; DB 1; Length 223;
Best Local Similarity 26.5%; Pred. No. 0.022;
Matches 48; Conservative 22; Mismatches 78; Indels 33; Gaps 6;

QY 2 AWM-----LLILIMVHFGSCALWVSQPPEIRTLLEGSAFLPCSFNASQGRLAIG 51
Dy 11 AWLELTSTRTALFSLFLPVSFGMHVAQPAVVLANSRGVASFVCEY--GSAGKAAEV 69

QY 52 SVTWRDEVVPGKEVRNGTPEFGRPLAPLASSRFLDHOAEHLHVRDVRGH 102
Dy 70 RVTVLRAGSQMTEVCAATYVDELT-----FLDSTCTGTSTENKVLTTQGLRAVD 123

QY 103 ASIYVCRVFL--GLGVGTGNGTRLVVKEHPOLGAGTVLLLR-----GFYAVSFLSVA 155
Dy 124 TGLYICKVELLYPPPYVVGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLITA 183

QY 156 V 156
Dy 184 V 184

RESULT 6
CTL4_RABIT STANDARD; PRT; 223 AA.
ID CTL4_RABIT AC P42072;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-
DE associated antigen 4) (CTLA-4).
GN CTLA4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=B/J X CHB;EM;
RX MEDLINE=95369849; PubMed=7642234;
RA Isono T., Seto A.;
RT "Cloning and sequencing of the rabbit gene encoding T-cell
RT costimulatory molecules.";
RL Immunogenetics 42:217-220(1995).
CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1
CC (CD80) AND B7-2 (CD86) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D49844; BAA08644.1; -.
CC FIR; I46696; I46696.
CC HSP; P16410; IAH1.
CC InterPro; IPR008096; CTLA4.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC PRINTS; PR01720; CTLANTIGEN4.
CC SMART; SM00409; IG; 1.
CC DR PROSITE; PS50835; IG LIKE; 1.
CC Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 35
FT CHAIN 36 223
FT DOMAIN 38 161
FT TRANSMEM 162 187
FT POTENTIAL.
FT CYTOTOXIC T-LYMPHOCYTE PROTEIN 4.
FT EXTRACELLULAR (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).

```

```

FT DOMAIN          36 145 IG-LIKE V-TYPE.
FT DISULFID        58 129 BY SIMILARITY.
FT DISULFID        85 103 BY SIMILARITY.
FT CARBOHYD        113 113 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD        145 145 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE        223 AA; 24655 MW; 85A9269793B88160 CRC64;

Query Match
Best Local Similarity 28.0%; DB 1; Length 223;
Matches 45; Conservative 22; Mismatches 83; Indels 11; Gaps 4;

QY 6 LLILIMVHPSGKALVWSPQPEIRTELGSSAFSCSNASQRLAIGSVTWFRDEVVPGKE 65
Db 25 LFSLSFLPVFSKALHVSQPAVLASSRGVASFVCEY-ASSHKATEYRTVTVLRQANSQMT 83
QY 66 VRNGTPEFRGRGLAPLASSR---FLRDHQAEHLHLDVRGHDSIYVCRVEVL---GLGVGT 119
Db 84 VCAMTYTVENELTIDFDTGTSIGSHGNKVNLTIGLSAMDTGLYCKVELMYPFVYVGM 143
QY 120 GNGTRLVVEKHPOLGAGTVLLRA----GFYAVSFUSVAV 156
Db 144 GNGQIYVIBPEPCPDSDFLLWILAAISSGLFFYSPLITAV 184

RESULT 7
SN_MOUSE
ID SN_MOUSE STANDARD; PRT; 1694 AA.
AC Q62230; Q55216; Q62228; Q62229;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-1) (Sheep erythrocyte receptor) (SER).
DE SN OR SA
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Macrophage;
RX MEDLINE=95009950; PubMed=792291;
RA Crocker P.R., Mucklow S., Boukson V., McWilliam A., Willis A.C.,
RA Gordon S., Milon G., Keim S., Bradfield P.;
RT "Sialoadhesin, a macrophage sialic acid binding receptor for
RT haemopoietic cells with 17 immunoglobulin-like domains.";
RL EMO J. 13:4490-4503(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=38051930; PubMed=9383289;
RA Mucklow S., Gordon S., Crocker P.R.;
RT "Characterization of the mouse sialoadhesin gene, Sn.";
RL Mamm. Genome 8:934-937(1997).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=91266893; PubMed=2050106;
RA Crocker P.R.;
RT "Purification and properties of sialoadhesin, a sialic acid-binding
RT receptor of murine tissue macrophages.";
RL EMO J. 10:1661-1669(1991).
RN [4]
RP SIALIC ACID BINDING.
RX MEDLINE=95179521; PubMed=7533044;
RA Keim S., Pelz A., Schauer R., Filbin M.T., Tang S., de Bellard M.E.,
RA Schnaar R.L., Mahoney J.A., Hartnell A., Bradfield P., Crocker P.R.;
RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new
RT family of sialic acid-dependent adhesion molecules of the
RT immunoglobulin superfamily.";
RL Curr. Biol. 4:965-972(1994).
RN [5]
RP BINDING TO SPN.
RX MEDLINE=21136329; PubMed=11238599;

```

```

RA van den Berg T.K., Nath D., Ziltener H.J., Vestweber D., Fukuda M.,
RA van Die I., Crocker P.R.;
RT "CD43 functions as a T cell counterreceptor for the macrophage
RT adhesion receptor sialoadhesin (Siglec-1).";
RL J. Immunol. 166:3637-3640(2001).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 20-138.
RX MEDLINE=93123385; PubMed=9660955;
RA May A.P., Robinson R.C., Vinson M., Crocker P.R., Jones E.Y.;
RT "Crystal structure of the N-terminal domain of sialoadhesin in
RT complex with 3' sialyllactose at 1.85 A resolution.";
RL Mol. Cell 1:719-728(1998).
RN [7]
RP STRUCTURE OF 20-138 BY NMR, AND MUTAGENESIS OF TRP-21 AND ARG-116.
RX MEDLINE=99321481; PubMed=10393093;
RA Crocker P.R., Vinson M., Kelm S., Drickamer K.;
RT "Molecular analysis of sialoside binding to sialoadhesin by NMR and
RT site-directed mutagenesis";
RL Biochem. J. 341:355-361(1999).
CC -!- FUNCTION: Macrophage-restricted adhesion molecule that mediates
CC sialic-acid dependent binding to lymphocytes, including
CC granulocytes, monocytes, natural killer cells, B-cells and CD8 T-
CC cells (By similarity). Preferentially binds to alpha2,3-linked
CC sialic acid. Binds to SPN/CD43 on T-cells. May play a role in
CC hemopoiesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and
CC soluble (isoforms 2 and 3).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q62230-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62230-2; Sequence=VSP_002573, VSP_002574;
CC Name=3;
CC IsoId=Q62230-3; Sequence=VSP_002575, VSP_002576;
CC -!- TISSUE SPECIFICITY: Expressed by macrophages in various tissues.
CC Highest expression in spleen and lymph node with lower amounts in
CC lung, liver, bone marrow, heart and skin. No expression in thymus,
CC kidney, brain or small intestine.
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC (sialic acid binding Ig-like lectin) family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z36293; CAA85290.1; -
CC EMBL; Z36233; CAA85268.1; -
CC EMBL; Z36234; CAA85269.1; -
CC EMBL; U92842; AAB95641.1; -
CC EMBL; U92833; AAB95641.1; JOINED.
CC EMBL; U92834; AAB95641.1; JOINED.
CC EMBL; U92836; AAB95641.1; JOINED.
CC EMBL; U92837; AAB95641.1; JOINED.
CC EMBL; U92838; AAB95641.1; JOINED.
CC EMBL; U92839; AAB95641.1; JOINED.
CC EMBL; U92840; AAB95641.1; JOINED.
CC EMBL; U92841; AAB95641.1; JOINED.
CC FIR; S50065; S50065.
CC PDB; 1QFO; 16-APR-99.
CC PDB; 1QFP; 16-APR-99.
CC MGD; MGI:99668; Sn.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0005529; F:sugar binding; ISS.
CC GO; GO:0016337; P:cell-cell adhesion; ISS.
CC GO; GO:0007160; P:cell-matrix adhesion; ISS.
CC GO; GO:0006954; P:inflammatory response; ISS.

```


to as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif is involved in downmodulation of cellular responses. The phosphorylated ITIM motif binds to the SH2 domain of PTEN6/SHP-1 and p185^{CD119}/SHP-2

PTM: Phosphorylated on tyrosine residues.
SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC (sialic acid binding Ig-like lectin) family.
SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.

: SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way misrepresented and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/send an email to license@isb-sib.ch>).

AF317298; AAG38598.1; -;
AF329269; AAK49917.1; -;
MGI:1932475; Siglec5.
GO:0004872; F:receptor activity; IDA.
GO:0005057; F:receptor signaling protein activity; IPI.
GO:0007155; P:cell adhesion; IDA.
IPR007110; Ig-like.
IPR003598; Ig_C2.
IPR003006; Ig_MHC.
PF00047; Ig; 2.
SM00408; IgC2; 1.
PS50835; IG_LIKE; 3.
PS02960; IG_MHC; 1.
adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
immunoglobulin domain; Repeat; Phosphorylation.

1	RES	455	POTENTIAL.
19	OHYD	467	SIALIC ACID BINDING IG-LIKE LECTIN 5.
19	OHYD	353	EXTRACELLULAR (POTENTIAL).
354	MSMEM	374	POTENTIAL.
375	MAIN	467	CYTOPLASMIC (POTENTIAL).
21	MAIN	141	IG-LIKE V-TYPE.
152	MAIN	239	IG-LIKE C2-TYPE 1.
239	MAIN	342	IG-LIKE C2-TYPE 2.
430	MAIN	435	ITIM MOTIF.
453	RES	458	SLAM-LIKE MOTIF.
40	LFID	176	BY SIMILARITY.
45	LFID	108	BY SIMILARITY.
170	LFID	219	BY SIMILARITY.
278	LFID	323	BY SIMILARITY.
432	RES	432	PHOSPHORYLATION (INVOLVED IN BINDING TO PTPN6 AND PTPN11) (PROBABLE).
455	RES	455	PHOSPHORYLATION (INVOLVED IN BINDING TO PTPN6) (PROBABLE).
46	OHYD	46	N-LINKED (GLCNAC. .) (POTENTIAL).
167	OHYD	167	N-LINKED (GLCNAC. .) (POTENTIAL).
197	OHYD	197	N-LINKED (GLCNAC. .) (POTENTIAL).
216	OHYD	216	N-LINKED (GLCNAC. .) (POTENTIAL).
227	OHYD	227	N-LINKED (GLCNAC. .) (POTENTIAL).
237	OHYD	237	N-LINKED (GLCNAC. .) (POTENTIAL).
244	OHYD	244	N-LINKED (GLCNAC. .) (POTENTIAL).
262	OHYD	262	N-LINKED (GLCNAC. .) (POTENTIAL).
287	OHYD	287	N-LINKED (GLCNAC. .) (POTENTIAL).
294	OHYD	294	N-LINKED (GLCNAC. .) (POTENTIAL).
432	AGEN	432	Y-F: ABOLISHES BINDING TO PTPN6 AND PTPN11.
455	AGEN	455	Y-F: REDUCES BINDING TO PTPN6.
2	YLIT	2	L -> M (IN REF. 2).
24	YLIT	24	V -> G (IN REF. 2).
102	YLIT	102	D -> E (IN REF. 2).
234	YLIT	234	T -> I (IN REF. 2).
467 AA;	ENCE	51889 MW;	54F03E98950D2DEE CRC64;

Query Match 9.7%; Score 99; DB 1; Length 467;
Best Local Similarity 26.0%; Pred. No. 0.12;
Matches 34; Conservative 21; Mismatches 50; Indels 26; Gaps 4;

Qy 4 MLLLLIMVHPGSCALWVSPPEIRTL-----EGSAFLPCSFNAGQRIATGSVT 54
| | | | : : : : | | | | : : : : | | | | : : : :
Db 1 MLLLLLLGLWGIGVEGQNPFVFILINVERKVVVGSLCVLPFCNFYSYLKKLTLTDWS 60
| | | | : : : : | | | | : : : : | | | | : : : :
Qy 55 -----WFRDE-----VVPKGVRNGTPEPRGSLAPLASSRFLHDHQAELHIDVRG 100
| | | | : : | | : : | | : : | | : : | | : :
Db 61 DPVHGFWYREGTRDRKDSIVATNNP-IRKAVKEFTRNEFFLLGDP--WRNDCSLNIREIK 117
| | | | : : | | : : | | : : | | : : | | : :
Qy 101 HDASIVCRVE 111
| | | | : : | | : : | | : : | | : :
Db 118 KDAGLYFFRL 128
| | | | : : | | : : | | : : | | : :

RESULT 9
ID BUTY_HUMAN STANDARD; PRT: 526 AA.
AC Q13410;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
De Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).
GN BTNL1 OR BTN
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxId=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=96201696; PubMed=8611614;
RA Taylor M.R., Peterson J.A., Ceriani R.L., Couto J.R.;
RT "Cloning and sequence analysis of human butyrophilin reveals a
RT potential receptor function";
RL Biochim. Biophys. Acta 1306:1-4(1996).
CC -I- FUNCTION: May function in the secretion of milk-fat droplets. It
CC may act as a specific membrane-associated receptor for the
CC association of cytoplasmic droplets with the apical plasma
CC membrane (By similarity).
CC -I- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase
CC (By similarity).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC family.
CC -I- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U39576; AAC50489.1; -.
CC PIR; S70587; S70587.
CC Genew; HGNC:1135; BTNL1A1.
CC MIT; G01610; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004872; F:receptor activity; TAS.
CC InterPro; IPRO01870; B302
CC InterPro; IPRO07110; Ig-like.
CC InterPro; IPRO03396; Ig.v.
CC InterPro; IPRO06574; PRY.
CC InterPro; IPRO03877; SPRY_receptor.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00622; SPRY; 1.
CC SMART; SM00406; IGV; 1.
CC SMART; SM00589; PRY; 1.
CC SMART; SM00449; SPRY; 1.
CC DR PROSITE; PS50835; IG_LIKE; 2.
CC TRANSmembrane; Glycoprotein; IMMUNOGLOBULIN domain; Signal; Repeat.
KW SIGNAL 1 26
FT

```

FT CHAIN 27 526 BUTYROPHILIN.
FT DOMAIN 27 242 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 243 269 POTENTIAL.
FT DOMAIN 270 526 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 138 IG-LIKE V-TYPE 1.
FT DOMAIN 148 234 IG-LIKE V-TYPE 2.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 526 AA; 59004 MW; E9EAC0CF8DAF94D5 CRC64;

Query Match 9.7%; Score 98.5; DB 1; Length 526;
Best Local Similarity 37.6%; Pred. No. 0.16;
Matches 44; Conservative 8; Mismatches 48; Indels 17; Gaps 8;

QY 4 MLLLLIMVHP-GSCALWVSQPE-IRTEGSAFLPC--SFNASQGRLAIGSVTWFRDE 59
DB 13 LTLILLQLPKLSDAPFDVIGPEPILAVVGDAELPCRLSPNASEHL---ELWFRKK 69
QY 60 VVPGKEV-RNG-----TPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVC 108
DB 70 VSPAVLHRDGRQEAQEPYRGR-ATLVQGIAGRVA-LRIRGRVSDDGGYTC 124

RESULT 10
MOG RAT
ID MOG RAT STANDARD; PRT; 245 AA.
AC Q63345;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Matthieu J.-M.;
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily.";
RL J. Neurosci. Res. 33:177-187(1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nusbaum J.-L., Roussel G.,
RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.P.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).
RN [3]
RP STRUCTURE BY NMR OF 62-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouze-Abou S., Wilson J.C., Bernard C.C.A., von Itzstein M.;
RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55.";
RL Eur. J. Biochem. 246:59-70(1997).
CC -1- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -1- SUBUNIT: May form homodimers.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC ACTIVE MYELINATION.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG

```

```

CC family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M99485; AAA41628.1; -
CC EMBL; L21995; AAF74786.1; -
CC PIR; B47712; B47712.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; Ig; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
CC Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
KW SIGNAL 1 27
FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 156 176 POTENTIAL.
FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 229 POTENTIAL.
FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 30 139 IG-LIKE.
FT DISULFID 51 125 POTENTIAL.
FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 245 AA; 27881 MW; C97F8AD60D6A32B4 CRC64;

Query Match 9.8%; Score 98; DB 1; Length 245;
Best Local Similarity 32.5%; Pred. No. 0.079;
Matches 39; Conservative 14; Mismatches 47; Indels 20; Gaps 7;

QY 1 MAMMLLILIMVHPGSCALWVSQPEIRTEGSAFLPCSFNASQGRLAIG-SVTWFRDE 59
DB 14 LSLULLLQLSRYAGQPRV-IGFCHPIRALVGDEALPC--RISPGKNATGMEVGHYRS- 69
QY 60 VVPGKEV-----RNG-----TPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVC 108
DB 70 --PFSRYVHLRYNGKQDAEAPEYRGRTELLKES--IGEGKVALRIQNVRFSDEGGYTC 125

RESULT 11
MOG_MOUSE
ID MOG_MOUSE STANDARD; PRT; 246 AA.
AC Q61885; P70364; Q62003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=95130110; PubMed=7829100;
RA Daubas P., Pham-Dinh D., Dautigny A.;
RT "Structure and polymorphism of the mouse myelin/oligodendrocyte
RT glycoprotein gene.";
RL Genomics 23:36-41(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gardiner M.V., Matthieu J.M.;
RT "Murine and human MOG are highly conserved: cDNA analysis.";
RL Trans. Am. Soc. Neurochem. 24:234-234(1993).

```


QY 108 CR-VEVLGLGVGTGNGTRLV 127
 Db 110 CQSFNTNQG-VFGGTKLTIV 129

RESULT 13
 BUTY_MOUSE
 ID BUTY_MOUSE STANDARD; PRT; 524 AA.
 AC Q62556; P97392;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).
 GN BTNL1 OR BTN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=Mammary gland;
 RX MEDLINE=97148936; PubMed=8995761;
 RA Ogg S.L., Komazagiri M.V.S., Mather I.H.;
 RT "Structural organization and mammary-specific expression of the
 RT butyrophilin gene";
 RL Mamm. Genome 7:900-905 (1996).
 RN [2]
 RP SEQUENCE OF 39-487 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=96125722; PubMed=8541302;
 RA Ishii T., Aoki N., Noda A., Adachi T., Nakamura R., Matsuda T.;
 RT "Carboxy-terminal cytoplasmic domain of mouse butyrophilin
 RT specifically associates with a 150-kDa protein of mammary epithelial
 RT cells and milk fat globule membrane";
 RL Biochim. Biophys. Acta 1245:283-292 (1995).
 CC -!- FUNCTION: May function in the secretion of milk-fat droplets. It
 CC may act as a specific membrane-associated receptor for the
 CC association of cytoplasmic droplets with the apical plasma
 CC membrane (By similarity).
 CC -!- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in mammary tissue and secreted in
 CC association with the milk-fat-globule membrane during lactation.
 CC -!- DEVELOPMENTAL STAGE: Expression increases during the last half of
 CC pregnancy and is maximal during lactation.
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
 CC family.
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U67065; AAB51034.1; -
 CC EMBL; S80642; AAB35893.1; -
 CC MGI; MG1:103118; Btntl1.
 CC InterPro; IPR001870; B302.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003596; IgV.
 CC InterPro; IPR006574; PRY.
 CC InterPro; IPR003877; SPRY_receptor.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF00622; SPRY; 1.
 CC SMART; SM00406; IG; 1.
 CC SMART; SM00589; PRY; 1.
 CC SMART; SM00449; SPRY; 1.
 CC PROSITE; PS50835; IG-LIKE; 2. Immunoglobulin domain; Signal; Repeat.
 CC Transmembrane; Glycoprotein; Immunoglobulin domain; Signal; Repeat.
 KW

FT SIGNAL 1 26
 FT CHAIN 27 524
 FT DOMAIN 27 524
 FT TRANSMEM 248 268
 FT DOMAIN 269 524
 FT DOMAIN 29 141
 FT DOMAIN 149 235
 FT CARBOHYD 56 56
 FT CARBOHYD 216 216
 FT CONFLICT 46 46
 FT CONFLICT 117 117
 FT CONFLICT 191 191
 FT CONFLICT 210 210
 FT CONFLICT 363 363
 FT CONFLICT 408 408
 FT CONFLICT 413 414
 FT CONFLICT 420 423
 FT CONFLICT 492 509
 FT
 SQ SEQUENCE 524 AA; 58406 MW; 333F4DEZC7704480 CRC64;

Query Match 9.3%; Score 95; DB 1; Length 524;
 Best Local Similarity 29.1%; Pred. No. 0.33;
 Matches 55; Conservative 17; Mismatches 51; Indels 66; Gaps 12;

QY 21 VSOPPE-IRTEGSSAFPCSF--NASQGLAIGSVTWFRD-----EYVPGKE 65
 Db 32 VTAQEPFVALVGSDAELTCGSPFNASSEYM---ELLNFRQTRSTAVLLYRDQGEQEQ 89
 QY 66 VRNGTPEFRGLAPLASSRFLDHQAEHLIRVGRHDASIVYCRVEVLGLGVGTGNGTRL 125
 Db 89 M-----TEYRGR-ATLTAGLL-DGRATLLIRDVRVSDQGEYRC----- 125
 QY 126 VVKEHPOLCAGVLLLRAGFYAVSFSLVAVGS-----TVYQKCHCHMGTHCHSSDG 179
 Db 126 -LFRNDNDFEEAAVYLKVA-----AVGSDPQISMTVQENG-----MELECTSS-- 168
 QY 180 PRGVIPEPR 188
 Db 169 --GWYEPQ 175

RESULT 14
 MOG_HUMAN
 ID -MOG_HUMAN STANDARD; PRT; 247 AA.
 AC Q16553; Q00713; Q00714; Q00715; Q13054; Q13055; Q14855; Q92891;
 AC Q92892; Q92893; Q92894; Q92895; Q93053; Q96KU9; Q96KV0; Q99605;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Myelin-oligodendrocyte glycoprotein precursor.
 GN MOG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=95310943; PubMed=7790876;
 RA Hilton A.A., Slavik A.J., Hilton D.J., Bernard C.C.A.;
 RT "Characterization of cDNA and genomic clones encoding human myelin
 RT oligodendrocyte glycoprotein";
 RL J. Neurochem. 65:309-318 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95054056; PubMed=7964757;
 RA Pham-Dinh D., Allinquant B., Ruberg M., della Gaspera B.,
 RA Nusebaum J.-L., Dautigny A.;
 RT "Characterization and expression of the cDNA coding for the human
 RT myelin/oligodendrocyte glycoprotein";
 RL J. Neurochem. 63:2353-2356 (1994).
 RN [3]

QY 72 --BFGELAPLASSRFLHDHQAELHIEDVGHDAIYVCEVEVLGLGVGTGNGTRLVVEX 129
Db 90 WQFGRGRVWVGDD--ISKHDSGIVIRNLDYINDGTFTCDVKRPPDVVGTSVDHLTVYD 146
QY 130 EHPQLGAGTV-----LLRAGFYAVSFL 152
Db 147 KIPVVGAGVWSGAIIGTFLGILLIIVGGLYLFYI 181

Search completed: February 26, 2004, 12:10:46
Job time : 18.1273 secs

RESULT 15
MYPO_HETP
ID MYPO_HETP STANDARD; PRT; 246 AA.
AC P20938;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral
protein) (MPP).
OS Heterodontus francisci (Horn shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Heterodontoides; Heterodontiformes;
OC Heterodontidae; Heterodontus.
CX NCBI_TaxID=7792;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90040744; PubMed=2478717;
RA Saavedra R.A., Fors L., Aebersold R.H., Arden B., Horvath S.,
RT "The myelin proteins of the shark brain are similar to the myelin
proteins of the mammalian peripheral nervous system."; J.
RL J. Mol. Evol. 29:149-156(1989).
CC -!- FUNCTION: Creation of an extracellular membrane face which guides
the wrapping process and ultimately compacts adjacent lamellae.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Found only in peripheral nervous system
Schwann cells.
CC -!- PTM: N-GLYCAN IS SULFATED (BY SIMILARITY).
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Belongs to the myelin P0 protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

CC EMBL; X16714; CAB37865.1; -.
CC PIR; A32999; A32999.
CC HSSP; P06907; INEU.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_V.
CC InterPro; IPR000920; Myelin_P0.
CC PRINTS; PR00213; MYELINP0.
CC SMART; SM00406; IGV; 1.
CC PROSITE; PS50835; IG LIKE; 1.
CC PROSITE; PS00568; MYELIN_P0; 1.
KW Myelin; Structural protein; Glycoprotein; Transmembrane;
KW Phosphorylation; Immunoglobulin domain; Signal.
FT SIGNAL 1 27
FT CHAIN 28 246 MYELIN P0 PROTEIN.
FT DOMAIN 28 150 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 151 178 POTENTIAL.
FT DOMAIN 179 246 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 143 IG-LIKE V-TYPE.
FT DISULFID 48 125 POTENTIAL.
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (COMPLEX) (BY
SIMILARITY).
FT
SQ SEQUENCE 246 AA; 27335 MW; A776A9ED9D430FA0 CRC64;

Query Match 9.1%; Score 93; DB 1; Length 246;
Best Local Similarity 23.9%; Pred. No. 0.23;
Matches 37; Conservative 26; Mismatches 66; Indels 26; Gaps 5;
QY 21 VSQPPRIETLGGSAFLPCSFNNSQGLAIGSVTW-FRD-----EVVPGKEVNGTFP---- 71
Db 30 VSTHNLKHTKVSQDVTLYCGFWSNEYVSDLTLSWFRPDNRDIIISIFHYGNVGYPIEK 89

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 28.7268 Seconds
(without alignments)
636.214 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 MAWLLLLILIMVHPGSCALW.....GTHCHSSDGRGVIPRCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	11.6	136	2 B45893	T-cell receptor al
2	117.5	11.5	132	2 A24402	T-cell receptor al
3	116	11.4	146	2 S26408	T-cell receptor be
4	115.5	11.3	232	2 S25756	Ig lambda chain -
5	111.5	10.9	226	2 A46477	membrane-bound imm
6	111	10.9	233	2 S25747	Ig lambda chain -
7	110	10.8	223	2 T09536	cytotoxic T-lympho
8	108.5	10.6	223	2 A29036	cytotoxic T-lympho
9	108	10.5	235	2 S25749	Ig lambda chain -
10	107.5	10.5	138	2 C27577	T-cell receptor al
11	106.5	10.4	131	2 E24092	T-cell receptor al
12	106.5	10.4	131	2 D24092	T-cell receptor al
13	106	10.4	132	1 RWSAV	T-cell receptor al
14	105	10.3	131	2 F45893	T-cell receptor al
15	104.5	10.2	132	2 S09713	Ig lambda chain V
16	103.5	10.1	223	2 I46896	CTLA-4 precursor
17	103	10.1	129	2 A42692	T-cell receptor al
18	102.5	10.0	110	2 E24092	T-cell receptor al
19	101.5	10.0	139	2 S36325	T-cell receptor de
20	101	9.9	120	2 I54487	T-cell receptor al
21	101	9.9	235	2 S14675	Ig lambda chain -
22	100.5	9.9	1694	2 S50065	sialoadhesin - mou
23	100	9.8	110	2 A24092	T-cell receptor al
24	99.5	9.8	271	2 A53268	T-cell receptor al
25	99	9.7	130	2 A31211	T-cell receptor al
26	98.5	9.7	128	2 S24319	Ig lambda chain pr
27	98.5	9.7	152	2 S21826	T-cell receptor be
28	98.5	9.7	236	2 S25746	Ig lambda chain -
29	98.5	9.7	247	2 A55717	myelin/oligodendro

30 98.5 9.7 526 2 S70587
31 96 9.4 131 1 I6HUBB
32 96 9.4 131 2 D24402
33 95.5 9.4 140 2 P0134
34 95.5 9.4 218 2 B47712
35 94.5 9.3 108 2 JH0342
36 94.5 9.3 186 2 S08614
37 93.5 9.2 247 2 S58394
38 93 9.1 117 2 I68824
39 93 9.1 132 2 D45893
40 93 9.1 246 1 A32999
41 93 9.1 267 1 RWS88
42 92.5 9.1 136 2 S42610
43 92.5 9.1 139 2 S36302
44 92.5 9.1 403 2 I52590
45 92 9.0 129 2 A30554

butyrophilin precu
Ig lambda chain pr
T-cell receptor al
Ig lambda chain pr
myelin/oligodendro
T-cell receptor al
cytotoxic T-lympho
myelin/oligodendro
T-cell receptor al
T-cell receptor al
myelin P0 protein
T-cell receptor al
ARMVlambda protein
T-cell receptor de
m33-B isoform - mo
Ig lambda chain pr

ALIGNMENTS

RESULT 1

B45893
T-cell receptor alpha chain precursor V region (BTA29) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C:Accession: B45893
R:Fishiguro, N.; Tanaka, A.; Shinagawa, M.
Immunogenetics 31, 57-60, 1990

A:Title: Sequence analysis of bovine T-cell receptor alpha chain.

A:Reference number: A45893; MUID:90129157; PMID:2137108

A:Accession: B45893

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-136 <ISR>

A:Cross-references: GB:D90011; NID:g217610; PIDN:BAA14061.1; PID:g217611

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: T-cell receptor

Query Match 11.6%; Score 118; DB 2; Length 136;
Best Local Similarity 27.9%; Pred. No. 0.0016;
Matches 41; Conservative 22; Mismatches 42; Indels 42; Gaps 9;

QY 4 MLLLLIMV-----HPGSCALWVSOPPEIRTLGSGSAFLPCFSNQGRLAIGSVTFW 56

Db 8 LLVILLASVDLEQPAEHS-----AEPASLPVEGAAASLGCTYSNSNIY----FTWY 57

QY 57 RDEVVPGKEVRNGTPEF-----RGELAPLASRFLHDHQAELHIRDVRGHDASI 105

Db 58 RQ--YPGK-----GPEFLQVYANNKEGKFT--AQSNTKNK--VSLIRDSPEPSDSAT 107

QY 106 YVCRVEVLGLGVT-----GNGTRLWVE 128

Db 108 YLCVAVDTTSTAGTKLTFEGTRLIIVK 134

RESULT 2

A24402
T-cell receptor alpha chain precursor V region (C5) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 30-May-1997

C:Accession: A24402

R:Becker, D.M.; Patten, P.; Chien, Y.; Yokota, T.; Eshhar, Z.; Giedlin, M.; Gascoigne,

Nature 317, 430-434, 1985

A:Title: Variability and repertoire size of T-cell receptor V-alpha gene segments.

A:Reference number: A33368; MUID:86014379; PMID:2995827

A:Accession: A24402

A:Molecule type: mRNA

A:Residues: 1-132 <BEC>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: T-cell receptor

Query Match 11.5%; Score 117.5; DB 2; Length 132;

```

Best Local Similarity 30.8%; Pred. No. 0.0018;
Matches 44; Conservative 17; Mismatches 47; Indels 35; Gaps 9;

Qy 12 VHGSCALW-----VSPPPE-IRTEGSSAFPCSFNASGRLAIGSVTWFRD 59
Db 3 LHVSLVFLW-LQGGVSSQKQVQSPESLIYPEGAMVSLNCSFSDS-----ASQSIMYQQH 58
Qy 60 VVQKQEVK-----NCTPFRGRPLAPLASSRFLHDHQAELHIRDVGRGHDASIYVCRVEV 112
Db 59 --FGKGPALISFSGNKK-EGRLTYLNRASH--VSLHKDQSPDSAVYLCAVR 112
Qy 113 LGLGVGT---GNGRLVVEKEHP 132
Db 113 SGANTGKLTGCHGTILRV---HP 132

RESULT 3
S26408
T-cell receptor beta chain V region (5.5) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26408
R: Bowman, S.J.; Lanchbury, J.S.
submitted to the EMBL Data Library, September 1992
A:Description: T cell receptor beta chain sequences from patients with rheumatoid arthritis
A:Reference number: S26408
A:Accession: S26408
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <BOW>
A:Cross-references: EMBL:X68527; NID:g36172; PIDN:CAA48540.1; PID:g36173
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:35-112/Domain: immunoglobulin homology <IMW>

Query Match 11.4%; Score 116; DB 2; Length 146;
Best Local Similarity 26.5%; Pred. No. 0.0027;
Matches 39; Conservative 26; Mismatches 52; Indels 30; Gaps 7;

Qy 1 MAMWLLILLIMVHGSCALWVSQPE--IRTEGSSAFPCSFNASGRLAIGSVTWFRD 59
Db 6 LCWLLCLL---GAGSVETGVTGFTLTKT-RGQVTLRCSSQSGH-----NTVSWYQQ 56
Qy 59 EWVFG-----KEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVGRGHDASIYVCRV 110
Db 57 ALGQGPQFIFYPYREENG-----RGNFPFSGLOFPNYSSELNVNALELDDSAVLICAS 112

Qy 111 EVLGLGVGT-----GNGRLVVEKE 130
Db 113 SFKGLGLPSRGYEQYFGFGRTLTVTED 139

RESULT 4
S25756
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25756
R: Combratio, G.; Klobbeck, H.G.
Eur. J. Immunol. 21, 1513-1522, 1991
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25756
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-232 <COM>
A:Cross-references: EMBL:X57821; NID:g33741; PIDN:CAA40958.1; PID:g33742
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:147-215/Domain: immunoglobulin homology <IMW>

Query Match 11.3%; Score 115.5; DB 2; Length 232;
Best Local Similarity 26.3%; Pred. No. 0.0049;

```

```

Matches 46; Conservative 24; Mismatches 62; Indels 43; Gaps 8;

Qy 1 MAMWLLILLIMVHGSCALWVSQPE--IRTEGSSAFPCSFNASGRLAIGSVTWFRD 58
Db 1 MAMWLLILLIMVHGSCALWVSQPE--IRTEGSSAFPCSFNASGRLAIGSVTWFRD 56
Qy 59 EWVFG-----KEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVGRGHDASIYV 107
Db 57 K--FQGPVAVVYDSDRPSGIPERFSGNS-----GNTATLTISRVEAGDEADYY 105
Qy 108 CRVEVLGCVGTGNGTRLV-----EKEHFGAGTIVLLIRAGFY 147
Db 106 CQVMDSSDVVFGGTKTLTVLQPKAAPSVTLFPPSSEELQANKATLVCLISDFY 160

RESULT 5
A46477
membrane-bound immunoglobulin Ig-alpha chain precursor - human
N:Alternate names: B-cell antigen receptor complex alpha chain CD79a; IgM-alpha; immunc
C:Species: Homo sapiens (man)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: I54539; A46477; A49135; A46479; I54496; I57851; S51113
R: Hashimoto, S.; Mohrenweiser, H.W.; Gregersen, P.K.; Chiorazzi, N.
Immunogenetics 40, 287-295, 1994
A:Title: Chromosomal localization, genomic structure, and allelic polymorphism of the I
A:Reference number: I54539; MUID:94364636; PMID:7916003
A:Accession: I54539
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <HAS1>
A:Cross-references: GB:L32754; NID:g600875; PIDN:AAA57274.1; PID:g602606
R: Yu, L.M.; Chang, T.W.
J. Immunol. 148, 633-637, 1992
A:Title: Human mb-1 gene: complete cDNA sequence and its expression in B cells bearing
A:Reference number: A46477; MUID:92105765; PMID:1729378
A:Accession: A46477
A:Molecule type: mRNA
A:Residues: 1-226 <YU>
A:Cross-references: GB:S75217; NID:g241773; PIDN:AA20812.1; PID:g241774
A:Note: sequence extracted from NCBI backbone (NCBIN:75217, NCBI:P:75218)
R: Leduc, I.; Preud'homme, J.L.; Cogne, M.
Clin. Exp. Immunol. 90, 141-146, 1992
A:Title: Structure and expression of the mb-1 transcript in human lymphoid cells.
A:Reference number: A49135; MUID:93009083; PMID:1395095
A:Accession: A49135
A:Molecule type: mRNA
A:Residues: 1-226 <LEB>
A:Cross-references: GB:S46706; NID:g257140; PIDN:AA23558.1; PID:g257141
A:Experimental source: Ly66 cell line
A:Note: sequence extracted from NCBI backbone (NCBIN:115213, NCBI:P:115214)
R: Ha, H.J.; Kubagawa, H.; Burrows, P.D.
J. Immunol. 148, 1526-1531, 1992
A:Title: Molecular cloning and expression pattern of a human gene homologous to the mu
A:Reference number: A46479; MUID:92168394; PMID:1538135
A:Accession: A46479
A:Molecule type: mRNA
A:Residues: 1-226 <HA1>
A:Cross-references: GB:U05259; NID:g452561; PIDN:AAA20495.1; PID:g521105; N
R: Flawinkel, H.; Reth, M.
Immunogenetics 36, 268-269, 1992
A:Title: Molecular cloning of the Ig-alpha subunit of the human B-cell antigen recepto
A:Reference number: I54496; MUID:92347937; PMID:1639443
A:Accession: I54496
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-68, 1' 70-226 <FLA>
A:Cross-references: GB:M74721; NID:g337419; PIDN:AAA60270.1; PID:g337420
R: Hashimoto, S.; Chiorazzi, N.; Gregersen, P.K.
Mol. Immunol. 32, 651-659, 1995
A:Title: Alternative splicing of CD79a (Ig-alpha/mb-1) and CD79b (Ig-beta/B29) RNA tra
A:Reference number: I57851; MUID:95371688; PMID:7643857
A:Accession: I57851

```

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-88, 'E', 128-226 <HAS3>
A:Cross-references: GB:S79248; NID:91087006; PIDN:AA60653.1; PID:91087007
R:Koyama M., Nakamura, T.
Submitted to the EMBL data Library, December 1994
A:Reference number: S5112
A:Accession: S5113
A:Molecule type: mRNA
A:Residues: 1-88, 'E', 128-226 <KOV>
A:Cross-references: EMBL:X83540; NID:9620080; PIDN:CAA59523.1; PID:9620081
C:Genetics:
A:Gene: GDB:TGA; mb-1; CD79a
A:Cross-references: GDB:133778; OMIM:600352
A:Map position: 19q13.2-19q13.2
A:Introns: 27/1; 127/1; 166/3; 190/1
C:Complex: disulfide-linked heterodimer with Ig-beta; heterodimer associates non-covalently
C:Keywords: alternative splicing; disulfide bond; glycoprotein; immunoglobulin; surface
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-226/Product: membrane-bound immunoglobulin Ig-alpha chain #status predicted <MAT>
F:33-143/Domain: extracellular #status predicted <EXT>
F:144-165/Domain: transmembrane #status predicted <TM>
F:166-226/Domain: intracellular #status predicted <CVT>
F:57,63,73,89,97,112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 111.5; DB 2; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.011; Mismatches 24; Conservative 47; Indels 70; Gaps 8;
Matches 47; Conservative 24; Mismatches 24; Indels 70; Gaps 8;

QY 4 MLLILLIMVHPGSCALWVSQ-PPEIRTLGSSAFPCSFNASQGRLAGSVTWFRDEVVP 62
DB 18 LFLSNAVILPGCCQALWVHKVPSALVSLGSDAHPQCPHNSNN-ANVTWR-VLH 71
QY 63 GKEVRNGT-PFPRGLAPLASSRFLHDAHRLHVRDVRGHDSIYVCRVEVLGVGTG 120
DB 72 G-NYTPPFLGP-----GDPNGLLIQNKSHGGIYVCRVQEGNESYQOS 118
QY 121 NGTFLVKEHFQ---LGAGTV-LLLRAGFVAVSFLSVAGSTVYQ 163
DB 119 CGTYLVRGPPRPFLDMGEGTNRITAEGLILLFCVAVPGTLLFR 166

RESULT 6
S25747
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25747
R:Combarato, G.; Klobbeck, H.G.
Eur. J. Immunol. 21, 1513-1522, 1991
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25747
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-233 <COM>
A:Cross-references: EMBL:X57812; NID:933723; PIDN:CAA40949.1; PID:933724
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:148-216/Domain: immunoglobulin homology <IMV>

Query Match 10.9%; Score 111; DB 2; Length 233;
Best Local Similarity 28.0%; Pred. No. 0.012; Mismatches 24; Conservative 49; Indels 60; Gaps 10;
Matches 49; Conservative 24; Mismatches 24; Indels 60; Gaps 10;

QY 1 MAWMLLLIMVH-PGSCALWV-SOPPEIRTLGSSAFPCSFNASQGRLAGSVTWFR 57
DB 1 MANTVLLGLLHSGTGSVTSYVLTQPPSVSVAPGVAARITCG-----GINTASVHWYQ 55
QY 58 DE-----VPGKEVR-NGTPE-FRGLAPLASSRFLHDAHRLHVRDVRGHDSIYVC 108
DB 56 QKPGQAPVLVWYGDSPGIPERFSGNS-----GNTATLNISRVEAGDEAAAYC 106

QY 109 RV-EVLGLGVGTGNGTRLVV-----EKEHPQLGAGTVLLLRAGFY 147
DB 107 QWVDSDDHVVFGGKILTVLGGPKAPSVTLFPPPSSELOANKATVCLISDFY 161

RESULT 7
T09536
cytotoxic T-lymphocyte protein 4 - human
C:Species: Homo sapiens (man)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T09536
R:Harper, K.; Balzano, C.; Rouvier, E.; Mattei, M.G.; Luciani, M.F.; Golstein, P.
J. Immunol. 147, 1037-1044, 1991
A:Title: CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mou
A:Reference number: I49584; MUID:91318145; PMID:1713603
A:Accession: T09536
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-223 <HAR>
A:Cross-references: EMBL:L15006; NID:9291928; PIDN:AA59385.1; PID:9291929
C:Genetics:
A:Gene: CTLA4
A:Map position: 2q33
A:Superfamily: T-cell surface glycoprotein CD28; immunoglobulin homology
C:Keywords: T-cell; transmembrane protein

Query Match 10.8%; Score 110; DB 2; Length 223;
Best Local Similarity 28.0%; Pred. No. 0.015; Mismatches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;
Matches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;

QY 6 LLLILVHVGSC-ALWVSOPPEIRTLGSSAFPCSFNASQGRLAGSVTWFRDEVVP 64
DB 24 LLEFLFIPVFCAMHVAQPAVVLASSRGIAFSVCEY-ASPGKATEVRVTILRQADSQVT 82
QY 65 EVANGPEPRGRIAPLASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL-- 113
DB 83 EVCAATVTGNETL-----FLDDSICTGTSSGNQNLTIQGLRMDTGLYICKVELMYP 136
QY 114 -GLGVGTGNGTRLVVEKEHPQLGAGTVLLLR-----GFYAVSFLSVAV 156
DB 137 PPTLGGTGTQIYVIDPEPCPDSDFLMLAAVSSGLFFYSFLITAV 184

RESULT 8
A29063
Cytotoxic T-lymphocyte protein 4 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 05-Nov-1999
C:Accession: A29063; I49622
R:Brust, J.F.; Denizot, F.; Luciani, M.F.; Roux-Dosseto, M.; Suzan, M.; Mattei, M.G.;
Nature 328, 267-270, 1987
A:Title: A new member of the immunoglobulin superfamily--CTLA-4.
A:Reference number: A29063; MUID:8758259; PMID:3496540
A:Accession: A29063
A:Molecule type: mRNA
A:Residues: 1-223 <BRU>
A:Cross-references: GB:X05719; NID:950592; PIDN:CAA2191.1; PID:950593
C:Superfamily: immunoglobulin V region; immunoglobulin homology
R:Harper, K.; Balzano, C.; Rouvier, E.; Mattei, M.G.; Luciani, M.F.; Golstein, P.
J. Immunol. 147, 1037-1044, 1991
A:Title: CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mou
A:Reference number: I49584; MUID:91318145; PMID:1713603
A:Accession: I49622
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-36 <RES>
A:Cross-references: GB:M74362; NID:912833; PIDN:AAA37489.1; PID:9553903
C:Genetics:
A:Gene: CtlA-4
A:Map position: 1, band C
C:Superfamily: T-cell surface glycoprotein CD28; immunoglobulin homology
C:Keywords: transmembrane protein

Query Match 10.6%; Score 108.5; DB 2; Length 223;

```
Best Local Similarity 26.9%; Pred. No. 0.02;
Matches 45; Conservative 26; Mismatches 79; Indels 17; Gaps 6;

QY 3 WMLLLILIMVPGSCALWVSOPPIRLTEGSAFLPCSFNASQGRLAIGSVTWFRD---- 58
DB 22 FVALLTLFIFVFSEAIQVTPQSVVLASSHGVAFPCEYSPSHNTDEV-RVTVLRQTNQ 80
QY 59 --EVVPGKEVRNGTPFERGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVL--- 113
DB 81 MTEVCATTFTEKNTVGFLD--YFPCSGTF-NESRVNLTIOGLRAVDGTGLVCKVELMYP 137
QY 114 GLGVGTGNGTFLVKEKHPQLGAGTVL-----LIRAGFYAVSFVSAV 156
DB 138 PYFVGMGNGTQIYVIDPECPDSDFLILWILVAVSLGLFFSYFLVSAV 184

RESULT 9
S25749
Ig lambda chain - human
C:Species: Homo sapiens (man)
C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
R:Combiarto, G.; Klobbeck, H.G.
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25749
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-235 <COM>
A:Cross-references: EMBL:X57814; NID:933727; PIDN:CAA40951.1; PID:933728
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:150-218/Domain: immunoglobulin homology <IM>

Query Match 10.6%; Score 108; DB 2; Length 235;
Best Local Similarity 25.1%; Pred. No. 0.023;
Matches 45; Conservative 26; Mismatches 60; Indels 48; Gaps 10;

QY 1 MAMWMLLILIMVHFGS---CALWVSOPPEIRTEGSAFLPCSFNASQGRLAIGSVTWFR 57
DB 1 MAMWMLLILGLLAY-GSGVDGQTVTVTPQPSFMSFGGTVTUTCLSSGSGSTSY-SPSWYQ 58
QY 58 DEVVPGKEVR-----NGTPE-FRGLAPLASSRFLHDHQAELHIRDVRGHDAIY 106
DB 59 Q--TPGQAPRTLMYNTTRSSGVDPFSGSIL-----GNKSALTITGAQAEDES 107
QY 107 VCRVEVLGLVGT---GNGTFLV-----EKEHPQLGAGTVLLLRAGFY 147
DB 108 YC-----VLYMGRGIVFGGKTLTVLGQPKAAPSVTLFPPPSSEELQANKATLVCLISDFY 163

RESULT 10
C27577
T-cell receptor alpha chain V region (5/10-20D) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-May-1997
A:Accession: C27577; C27557
R:Iwamoto, A.; Ohashi, P.S.; Picher, H.; Walker, C.L.; Michalopoulos, E.E.; Rupp, F.; H
J. Exp. Med. 165, 591-600, 1987
A:Title: T cell receptor variable gene usage in a specific cytotoxic T cell response. P
A:Reference number: A27557; MUID:87139812; PMID:3493320
A:Accession: C27577
A:Molecule type: mRNA
A:Residues: 1-138 <IWA>
A:Accession: C27557
A:Molecule type: mRNA
A:Residues: 22-138 <IW2>
A:Cross-references: EMBL:X05733
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 10.5%; Score 107.5; DB 2; Length 138;
```

```
Best Local Similarity 31.9%; Pred. No. 0.015;
Matches 38; Conservative 17; Mismatches 45; Indels 19; Gaps 7;

QY 21 VSQPEE-IRTEGSAFLPCSFNASQGRLAIGSVTWFRDEVVPGKEVR-----NGTPE 72
DB 24 VQQSPESLIVPEGAMTSLNCTFSDASQY-----FMYRQH--SGKAPKALMSIFSNGEKE 77
QY 73 FRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVL-GLGVGTGNGTFLVVEKE 130
DB 78 -EGRFTIHLNKSLH---FSLHIRDSQPSDSALYLCAVSMGTGGYKVVFGSGTLLVSPD 132

RESULT 11
E24092
T-cell receptor alpha chain precursor V region (4.C3) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 21-Jan-2000
C:Accession: E24092
R:Pink, P.J.; Mattis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A:Title: Correlations between T-cell specificity and the structure of the antigen recep
A:Reference number: A93380; MUID:86230843; PMID:3012351
A:Accession: E24092
A:Molecule type: mRNA
A:Residues: 1-131 <FIN>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:21-131/Product: T-cell receptor alpha chain V region 4.C3 #status predicted <MAT>
F:36-109/Domain: immunoglobulin homology <IM>

Query Match 10.4%; Score 106.5; DB 2; Length 131;
Best Local Similarity 29.6%; Pred. No. 0.017;
Matches 37; Conservative 19; Mismatches 42; Indels 27; Gaps 7;

QY 21 VSQPEIRTL-EGSSAFPCSFNASQGRLAIGSVTWFR-----RDRVWPGKEVRNGTPEFRG 75
DB 24 VEQSPSALSLEHGTGSAALRCNFTTT-----MRAVQWRKNSRGLINFLYASGTKE-NG 77
QY 76 RLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGNGTFLVVEKEHPQLG 135
DB 78 RLKSAFDSK---BRYSTLHIRDAQLSDSGTYFCAE-----STSSQGLV-----FG 121
QY 136 AGTVL 140
DB 122 QGTVL 126

RESULT 12
D24092
T-cell receptor alpha chain precursor V region (B10) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1988 #sequence_revision 01-Dec-2000 #text_change 11-Jan-2002
C:Accession: D24092; S03507
R:Pink, P.J.; Mattis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A:Title: Correlations between T-cell specificity and the structure of the antigen recep
A:Reference number: A93380; MUID:86230843; PMID:3012351
A:Accession: D24092
A:Molecule type: mRNA
A:Residues: 1-131 <FIN>
R:Iwamoto, A.; Mjolsness, S.; Hood, L.
Nature 316, 832-836, 1985
A:Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.
A:Reference number: S03503; MUID:85296332; PMID:2993908
A:Accession: S03507
A:Molecule type: DNA
A:Residues: 111-130 <WIN>
A:Cross-references: EMBL:X03057; NID:954519; PIDN:CAA26864.1; PID:gl334131
A>Note: this sequence was determined from the germline gene
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:1-20/Domain: signal sequence #status predicted <SIG>
```

F;21-131/Product: T-cell receptor alpha chain V region B10 #status predicted <MAT>
F;36-109/Domain: immunoglobulin homology <IMM>

Query Match 10.4%; Score 106.5; DB 2; Length 131;
Best Local Similarity 29.6%; Pred. No. 0.017; Indels 27; Gaps 7;
Matches 37; Conservative 19; Mismatches 42; Indels 27; Gaps 7;
QY 21 VSOPPEIRTL-EGSSAFPCSFNASQGRLAIGSVTWTF-----RDEVVPGKEVRNGTPEFRG 75
Db 24 VEQSPSALSHEGTGSALRCNFTT-----MRAVQFRKNRSGSLNLFVLSAGTKE-NG 77
QY 76 RLAPLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGVGTGTGTLVVEKEHPQLG 135
Db 78 RLKSAFDK---ERYSTLHIDAQLESDGTGYFCAAE-----ATSSGQKLV-----FG 121
QY 136 AGTVL 140
Db 122 QGTVL 126

RESULT 13

RWMSAV
T-cell receptor alpha chain precursor V region (2B4) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 30-May-1997
C;Accession: A02015
R;Chien, Y.; Becker, D.M.; Lindsten, T.; Okamura, M.; Cohen, D.I.; Davis, M.M.
Nature 312, 31-35, 1984
A;Title: A third type of murine T-cell receptor gene.
A;Reference number: A93344; MUID:85036634; PMID:6548551
A;Accession: A02015
A;Molecule type: mRNA
A;Residues: 1-132 <CHI>
A;Experimental source: hybridoma 2B4, clone T11
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: glycoprotein; heterotetramer; receptor; T-cell
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-132/Product: T-cell receptor alpha chain V region (2B4) #status predicted <MAT>
F;21-113/Region: D segment
F;114-117/Region: D segment
F;118-132/Region: J segment
F;42/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 106; DB 1; Length 132;
Best Local Similarity 25.8%; Pred. No. 0.019; Indels 48; Gaps 9;
Matches 40; Conservative 22; Mismatches 45; Indels 48; Gaps 9;
QY 3 WMLLLILIMVH-PGSCALWY-----QPPEIRTEGSAFLPCSFNASQGRLAIGSVTW 55
Db 12 WLL-----NWNSQNVQOSPESLIVPEGATSLNCTFSDASQY-----FW 55
QY 56 PRDEVVPGKEVR-----NGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDASIVY 108
Db 56 YRQH--SGKAPKALMSIFSGEKE--SGRTIHLNKASLH---FSLHIRDSQPSDSALYIC 109
QY 109 RVEVLGVGTGTGTLVVEKEHPQIGAGTVLLLR 143
Db 110 AVTLYG-----GSGNKL-----FGTGTLLSVK 132

RESULT 14

F45893
T-cell receptor alpha chain precursor V region (BTA25) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000
C;Accession: F45893
R;Ishiguro, N.; Tanaka, A.; Shinagawa, M.
Immunogenetics 31, 57-60, 1990
A;Title: Sequence analysis of bovine T-cell receptor alpha chain.
A;Reference number: A45893; MUID:90129157; PMID:2137108
A;Accession: F45893
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-131 <ISH>
A;Cross-references: GB:D50015; NID:g217618; PIDN:BA14065.1; PID:g217619
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor
F;36-109/Domain: immunoglobulin homology <IMM>

Query Match 10.3%; Score 105; DB 2; Length 131;
Best Local Similarity 30.3%; Pred. No. 0.023; Indels 26; Gaps 8;
Matches 36; Conservative 23; Mismatches 34; Indels 26; Gaps 8;
QY 21 VSOPPEIRTL-EGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKE-----VNGTPEF 73
Db 24 VEQSPSLSLQEGANSTLRNFSDT-----VDSYQVFQON--PGGALTTLFFIASGTKK- 75
QY 74 RGLAPLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGVGT-----GNGTRLVV 127
Db 76 NERMSSTVNSK---ERYSTLHITASQLEDAATVLCVD---LSSGNQLVFGKGRRLAV 128

RESULT 15

S09713
IG lambda chain V region - human
C;Species: Homo sapiens (man)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C;Accession: S09713
R;Hughes-Jones, N.C.; Bye, J.M.; Beale, D.; Coadwell, J.
Biochem. J. 268, 135-140, 1990
A;Title: Nucleotide sequences and three-dimensional modelling of the VH and VL domains
A;Reference number: S09710; MUID:90262535; PMID:2111699
A;Accession: S09713
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-132 <HUG>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-110/Domain: immunoglobulin homology <IMM>

Query Match 10.2%; Score 104.5; DB 2; Length 132;
Best Local Similarity 29.8%; Pred. No. 0.026; Indels 25; Gaps 8;
Matches 42; Conservative 20; Mismatches 54; Indels 25; Gaps 8;
QY 1 MAMVLLILIMVH-PGSCALWY-SQPPEIRTEGSAFLPCSFNASQGRLAIGSVTWFRD 58
Db 1 MAMTVLLGLLGHCTGTSVTSVLTQPPSVSVPQQTARITC--GGTSNNIGRKSVMHYQQ 58
QY 59 E-----VVPCKEVR-NGTPE-FRGLAPLASSRFLHDHQAELHIRDVRGHDASIVYCR 109
Db 59 KPGQAPVLVYVGASDRPSGIPERFSGSNS-----GNTATLTISRVAAGDEADYYCQ 109
QY 110 V---EVILGLGVGTGNGTRLVV 127
Db 110 WVDSSSAHPGVFGGGLKLV 130

Search completed: February 26, 2004, 12:14:49
Job time : 30.7268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 117.931 Seconds
(without alignments)
455.215 Million cell updates/sec

Title: US-10-036-444-2
Perfect score: 1020
Sequence: 1 MAWMLLLILIMVHFGSCALW.....GTHCHSSDGRGVIPEPRCP 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 26Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	190	AAE02769	Aae02769 Human NKp
2	1018	99.8	190	AAY06401	Aay06401 Human B-c
3	876	85.9	201	AAY06403	Aay06403 Human B-c
4	859	84.2	177	AAY06402	Aay06402 Human B-c
5	713	69.9	135	AAE19109	Aae19109 Human NKp
6	713	69.9	369	AAE19110	Aae19110 Human NKp
7	632	62.0	120	AAE02771	Aae02771 Human NKp
8	204	20.0	33	AAE02773	Aae02773 Human NKp
9	123.5	12.1	246	AAE20273	Aae20273 Human lun
10	119.5	11.7	232	ABU08020	Abu08020 Monoclonal
11	118.5	11.6	139	AU75565	Au75565 Murine T
12	118.5	11.6	212	AAE60134	Aae60134 CTLA4 rec
13	118.5	11.6	212	AAE77642	Aar77642 Full leng
14	118.5	11.6	212	AAE43479	Aay43479 Amino aci
15	118.5	11.6	212	AAW81584	Aaw81584 Human CTL
16	118.5	11.6	212	ABW78106	Abb78106 Amino aci
17	118.5	11.6	212	AAU75124	Aau75124 Human CTL
18	118.5	11.6	212	AAU75132	Aau75132 Human CTL
19	118.5	11.6	212	ABP56715	Abp56715 Human CTL
20	118.5	11.6	212	ADD89013	Adg89013 CTLA4 rec
21	116.5	11.4	212	AAU75126	Aau75126 Human CTL
22	116	11.4	211	AAW87560	Aaw87560 Human CTL
23	115.5	11.3	212	AAU75127	Aau75127 Human CTL
24	115.5	11.3	212	AAU75131	Aau75131 Human CTL
25	115.5	11.3	212	AAU75130	Aau75130 Human CTL

26	114.5	11.2	212	5	AAU75133	Aau75133 Human CTL
27	112	11.0	222	6	ABP5948	Abp5948 Human CTL
28	112	11.0	223	3	AAU75129	Aay75129 Human CTL
29	112	11.0	223	5	AAU74509	Aau74509 Human cyt
30	112	11.0	223	5	AAU74508	Aau74508 Human cyt
31	112	11.0	235	4	AAU74508	Aau74508 Human cyt
32	111.5	10.9	212	2	AAU74508	Aau74508 Human cyt
33	111.5	10.9	212	5	AAU75129	Aau75129 Human CTL
34	111.5	10.9	226	2	AAU75129	Aau75129 Human CTL
35	111.5	10.9	226	3	AAU75129	Aau75129 Human CTL
36	111.5	10.9	226	6	ABG74347	Abg74347 Human Iga
37	111.5	10.9	226	7	ADD67577	Add67577 Human Ly6
38	110.5	10.8	223	5	ABB79935	Abb79935 Mouse CTL
39	110	10.8	223	2	AAU25111	Aau25111 Soluble h
40	110	10.8	223	4	AAU00687	Aau00687 Human CTL
41	110	10.8	223	5	ABB79934	Abb79934 Human CTL
42	110	10.8	223	7	ADA50039	Ada50039 Human cyt
43	110	10.8	223	7	ADC78847	Adc78847 Human PRO
44	110	10.8	223	7	ADD25540	Add25540 Binding d
45	110	10.8	235	4	AAG64476	Aag64476 Human typ

ALIGNMENTS

RESULT 1

AAE02769
ID AAE02769 standard; protein; 190 AA.

XX AC AAE02769;

XX DT 06-AUG-2001 (first entry)

XX DE Human NKp30 receptor.

XX KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;

XX KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

XX KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..18

XX FT Protein 19..190

XX FT Region 19..138

XX FT Modified-site 42

XX FT Modified-site 121

XX FT Region 139..157

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

XX FT Region 158..190

DR WPI; 2001-329221/34.
 DR N-PSDB; AAD06564.
 XX Novel compound, useful for detection and/or quantifying the presence of
 PT NK cells, comprises the amino acid sequences of the Nkp30 molecule.
 XX
 PS Claim 1; Fig 7B; 83pp; English.
 XX
 CC The invention relates to human Nkp30 receptor and its corresponding cDNA
 CC molecule which is involved in natural cytotoxicity mediated by natural
 CC killer (NK) cells and antibodies that identify the same. Nkp30 receptor
 CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively
 CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are
 CC useful for detecting and/or quantifying the presence of NK cells in a
 CC biological sample. The invention also provide kits for detecting and/or
 CC quantifying the presence of NK cells, for the selective removal of NK
 CC cells from a biological sample, for the positive and selective
 CC purification of NK cells from a biological sample and for the in vitro
 CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GvH) inhibition, stimulation of graft
 CC versus tumour (GvT) and especially graft versus leukaemia (GvL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. Nkp30 antibodies are useful for
 CC identifying Nkp30 natural ligands and allow assessment of the level of
 CC surface Nkp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence Nkp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is human Nkp30 receptor
 XX
 SQ Sequence 190 AA;

Query Match 100.0%; Score 1020; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 6.6e-92;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAWMLLLILMVHPGSCALWVSQPPETRTLEGSSAFPCSFNASQGRLAIGSVTFRDEV 60
 DB 1 MAWMLLLILMVHPGSCALWVSQPPETRTLEGSSAFPCSFNASQGRLAIGSVTFRDEV 60
 QY 61 VPGKEVRNGTPEFRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 DB 61 VPGKEVRNGTPEFRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 QY 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYQKCHCHMGTHCHSSDGP 180
 DB 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYQKCHCHMGTHCHSSDGP 180
 QY 181 RGVIPEPRCP 190
 DB 181 RGVIPEPRCP 190

RESULT 2
 AAY06401
 ID AAY06401 standard; protein; 190 AA.
 AC AAY06401;
 XX
 XX 20-SEP-1999 (first entry)
 DT Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 DE BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
 KW signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..12
 FT /note= "leader peptide"

FT Protein 13. .190
 FT /note= "mature protein"
 FT Modified-site 42
 FT /note= "N-glycosylated"
 FT Modified-site 68
 FT /note= "N-glycosylated"
 FT Modified-site 121
 FT /note= "N-glycosylated"
 FT Domain 139..162
 FT /note= "transmembrane domain"
 FT Peptide 166..190
 FT /note= "alternatively spliced C-terminal end"
 PN WO9923867-A2.
 XX 20-MAY-1999.
 XX 05-NOV-1998; 98WO-US023826.
 XX 07-NOV-1997; 97US-0064761P.
 XX (BIOJ) BIOGEN INC.
 XX Browning J;
 XX WPI; 1999-418423/35.
 XX N-PSDB; AAX59347.
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 PS Claim 2; Page 42; 43pp; English.

This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

Query Match 99.8%; Score 1018; DB 2; Length 190;
 Best Local Similarity 99.5%; Pred. No. 1e-91;
 Matches 189; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAWMLLLILMVHPGSCALWVSQPPETRTLEGSSAFPCSFNASQGRLAIGSVTFRDEV 60
 DB 1 MAWMLLLILMVHPGSCALWVSQPPETRTLEGSSAFPCSFNASQGRLAIGSVTFRDEV 60
 QY 61 VPGKEVRNGTPEFRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 DB 61 VPGKEVRNGTPEFRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 QY 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYQKCHCHMGTHCHSSDGP 180
 DB 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYQKCHCHMGTHCHSSDGP 180
 QY 181 RGVIPEPRCP 190
 DB 181 RGVIPEPRCP 190

RESULT 3
 AAY06403
 ID AAY06403 standard; protein; 201 AA.

XX AAY06403;
 AC 20-SEP-1999 (first entry)
 DT Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 DE MBOG; B-cell myelin oligodendrocyte glycoprotein; human;
 XX signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 KW OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..12
 FT /note= "leader peptide"
 FT Protein 13..201
 FT /note= "mature protein"
 FT Modified-site 42
 FT /note= "N-glycosylated"
 FT Modified-site 68
 FT /note= "N-glycosylated"
 FT Modified-site 121
 FT /note= "N-glycosylated"
 FT Domain 139..162
 FT /note= "transmembrane domain"
 FT Peptide 166..201
 FT /note= "alternatively spliced C-terminal end"
 XX
 XX WO9923867-A2.
 XX
 XX 20-MAY-1999.
 XX
 XX 05-NOV-1998; 98WO-US023826.
 XX
 XX 07-NOV-1997; 97US-0064761P.
 XX
 XX (BIOJ) BIOGEN INC.
 XX
 XX Browning J;
 XX
 XX WPI; 1999-418423/35.
 XX N-PSDB; AAX59349.
 XX
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 XX
 XX Claim 2; Page 43; 43pp; English.
 XX
 XX This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 XX Sequence 201 AA;
 XX
 XX Query Match 85.9%; Score 876; DB 2; Length 201;
 XX Best Local Similarity 89.2%; Pred. No. 9.9e-78;
 XX Matches 173; Conservative 2; Mismatches 7; Indels 12; Gaps 2;
 QY 1 MAWMLLLIIMVHPSGALWVSQPPPIRTLEGSSAFPCSFNAGSLAIGSVTFWRDEV 60
 DB 1 MAWMLLLIIMVHPSGALWVSQPPPIRTLEGSSAFPCSFNAGSLAIGSVTFWRDEV 60

QY 61 VPGKEVRNGTPEFRGRGLAPLASSRFLHDHQAELHIRDVRGHDSIYVCRVEVLGLGVGTG 120
 DB 61 VPGKEVRNGTPEFRGRGLAPLASSRFLHDHQAELHIRDVRGHDSIYVCRVEVLGLGVGTG 120
 QY 121 NGTRLVVEKEHPOLGAGTIVLLRAGFYAVSFLSVAGSTVYVYQKCHMGTHCHSDGP 180
 DB 121 NGTRLVVEKEHPOLGAGTIVLLRAGFYAVSFLSVAGSTVYVYQKCHMGTHCHSDGP 180
 QY 181 R---GVPEPRCP 190
 DB 173 RRQLPAVVPAPLPP 186
 RESULT 4
 AAY06402
 ID AAY06402 standard; protein; 177 AA.
 XX
 XX AC AAY06402;
 XX
 XX 20-SEP-1999 (first entry)
 DT Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 DE MBOG; B-cell myelin oligodendrocyte glycoprotein; human;
 XX signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 KW OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..12
 FT /note= "leader peptide"
 FT Protein 13..177
 FT /note= "mature protein"
 FT Modified-site 42
 FT /note= "N-glycosylated"
 FT Modified-site 68
 FT /note= "N-glycosylated"
 FT Modified-site 121
 FT /note= "N-glycosylated"
 FT Domain 139..162
 FT /note= "transmembrane domain"
 FT Peptide 166..177
 FT /note= "alternatively spliced C-terminal end"
 XX
 XX WO9923867-A2.
 XX
 XX 20-MAY-1999.
 XX
 XX 05-NOV-1998; 98WO-US023826.
 XX
 XX 07-NOV-1997; 97US-0064761P.
 XX
 XX (BIOJ) BIOGEN INC.
 XX
 XX Browning J;
 XX
 XX WPI; 1999-418423/35.
 XX N-PSDB; AAX59348.
 XX
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 XX
 XX Claim 2; Page 43; 43pp; English.
 XX
 XX This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 XX Sequence 201 AA;
 XX
 XX Query Match 85.9%; Score 876; DB 2; Length 201;
 XX Best Local Similarity 89.2%; Pred. No. 9.9e-78;
 XX Matches 173; Conservative 2; Mismatches 7; Indels 12; Gaps 2;
 QY 1 MAWMLLLIIMVHPSGALWVSQPPPIRTLEGSSAFPCSFNAGSLAIGSVTFWRDEV 60
 DB 1 MAWMLLLIIMVHPSGALWVSQPPPIRTLEGSSAFPCSFNAGSLAIGSVTFWRDEV 60

CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX

SQ Sequence 177 AA;
 Query Match 84.2%; Score 859; DB 2; Length 177;
 Best Local Similarity 100.0%; Pred. No. 3.9e-76;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGRLAIGSVTWRDEV 60
 DB 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGRLAIGSVTWRDEV 60
 QY 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 DB 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 QY 121 NGTRLVVEKEHPOLGAGTVLLRAGFYAVSFVSVAVGSTVYQYQK 165
 DB 121 NGTRLVVEKEHPOLGAGTVLLRAGFYAVSFVSVAVGSTVYQYQK 165

RESULT 5
 AAE19109
 ID AAE19109 standard; protein; 135 AA.
 XX
 AC AAE19109;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human NKp30 protein.
 XX
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; NKp30.
 XX
 OS Homo sapiens.
 OS
 PN WO200208287-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-IL000664.
 XX
 PR 20-JUL-2000; 2000IL-00137419.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA (UYNE) UNIV BEN-GURION NEGEV.
 XX
 PI Mandelboim O, Porgador A;
 XX
 DR WPI; 2002-195870/25.
 DR N-PSDB; AAD30466.
 XX
 XX New targeting complex capable of targeting an active substance to a
 PT target cell, comprising a target recognition segment and an active
 PT segment, useful for treating pathologies associated with viral infections
 PT or cancer.
 XX
 PS Example 1; Page 108; 113pp; English.

XX The invention relates to compositions and methods for the treatment and
 CC detection of a variety of viral infections, by using complex agents
 CC comprising the natural killer (NK) cells activating proteins, NKp46 and
 CC NKp44 and functional fragments thereof, linked to therapeutic or imaging
 CC agents. The complex is useful for treating pathologies associated with
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECHOV, MCM or herpes virus)
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
 CC the imaging and monitoring of cancer. The complex may also be used to

CC detect the presence of abnormal cells in a sample. The antibodies can be
 CC used to qualitatively or quantitatively detect the ligand for the
 CC complex. The present sequence is human NKp30 protein
 XX
 SQ Sequence 135 AA;

Query Match 69.9%; Score 713; DB 5; Length 135;
 Best Local Similarity 100.0%; Pred. No. 6.1e-62;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGRLAIGSVTWRDEV 60
 DB 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGRLAIGSVTWRDEV 60
 QY 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 DB 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
 QY 121 NGTRLVVEKEHPOLG 135
 DB 121 NGTRLVVEKEHPOLG 135

RESULT 6
 AAE19110
 ID AAE19110 standard; protein; 369 AA.
 XX
 AC AAE19110;
 XX
 DT 29-AUG-2003 (revised)
 DT 21-MAY-2002 (first entry)
 XX
 DE Human NKp30-IgG fusion protein.
 XX
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;
 KW immunoglobulin G; fusion protein.
 XX
 OS Homo sapiens.
 OS Chimeric.
 FH Key Location/Qualifiers
 FT Region 1..135
 FT /note= "Human NKp30"
 FT Region 136..369
 FT /note= "Human IgG"
 XX
 PN WO200208287-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-IL000664.
 XX
 PR 20-JUL-2000; 2000IL-00137419.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA (UYNE) UNIV BEN-GURION NEGEV.
 XX
 PI Mandelboim O, Porgador A;
 XX
 DR WPI; 2002-195870/25.
 DR N-PSDB; AAD30467.
 XX
 XX New targeting complex capable of targeting an active substance to a
 PT target cell, comprising a target recognition segment and an active
 PT segment, useful for treating pathologies associated with viral infections
 PT or cancer.
 XX
 PS Example 1; Page 108-110; 113pp; English.
 XX The invention relates to compositions and methods for the treatment and
 CC detection of a variety of viral infections, by using complex agents
 CC comprising the natural killer (NK) cells activating proteins, NKp46 and
 CC NKp44 and functional fragments thereof, linked to therapeutic or imaging
 CC agents. The complex is useful for treating pathologies associated with
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECHOV, MCM or herpes virus)
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
 CC the imaging and monitoring of cancer. The complex may also be used to

comprising the natural killer (NK) cells activating proteins, Nkp46 and Nkp44 and functional fragments thereof, linked to therapeutic or imaging agents. The complex is useful for treating pathologies associated with viral infections (e.g. infections caused by influenza virus, HIV, Epstein-Barr virus, cytomegalovirus, vaccinia virus, HCMV, MVM or herpes virus) and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for the imaging and monitoring of cancer. The complex may also be used to detect the presence of abnormal cells in a sample. The antibodies can be used to qualitatively or quantitatively detect the ligand for the NKp30 complex. The present sequence is human NKp30- immunoglobulin G (IgG) Fc region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 369 AA;

Query Match 69.9%; Score 713; DB 5; Length 369;
Best Local Similarity 100.0%; Pred. No. 2.2e-61;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAWMLLLILIMVHPGSCALWVSOPPEIRTLGSSAPLPCSFNASQGRLAIGSVTFRDEV 60
DB 1 MAWMLLLILIMVHPGSCALWVSOPPEIRTLGSSAPLPCSFNASQGRLAIGSVTFRDEV 60

QY 61 VPGKEVRNGTPEFRGRFLAPLASSRFLHDHQAELHIRDVRGHDASTYVCRVEVLGLGVGTG 120
DB 61 VPGKEVRNGTPEFRGRFLAPLASSRFLHDHQAELHIRDVRGHDASTYVCRVEVLGLGVGTG 120

QY 121 NGTRLVVEKEHPOLG 135
DB 121 NGTRLVVEKEHPOLG 135

RESULT 7
AAE02771
ID AAE02771 standard; protein; 120 AA.

XX AAE02771;
XX AC AAE02771;
XX DT 06-AUG-2001 (first entry)
XX DE Human NKp30 receptor extracellular region sequence.
XX DE Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;
XX KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
XX KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
XX KW therapy; extracellular region.
XX OS Homo sapiens.
XX PN WO200136630-A2.
XX PD 25-MAY-2001.
XX PF 15-NOV-2000; 2000WO-EP011697.
XX PR 15-NOV-1999; 99CA-02288307.
XX PR 15-NOV-1999; 99US-00440514.
XX XX (INNA-) INNATE PHARMA SAS.
XX PA (UYGE-) UNIV GENOVA.
XX PI Moretta A, Bottino C, Biassoni R;
XX XX WPI; 2001-329221/34.
XX DR Novel compound, useful for detection and/or quantifying the presence of
XX PT NK cells, comprises the amino acid sequences of the NKp30 molecule.
XX XX Claim 1; Fig 7B; 83pp; English.

XX The invention relates to human NKp30 receptor and its corresponding cDNA
XX molecule which is involved in natural cytotoxicity mediated by natural
XX killer (NK) cells and antibodies that identify the same. NKp30 receptor
XX is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively

expressed on the surface of human mature NK cells. NKp30 and its cDNA are useful for detecting and/or quantifying the presence of NK cells in a biological sample. The invention also provides kits for detecting and/or quantifying the presence of NK cells, for the selective removal of NK cells from a biological sample, for the positive and selective purification of NK cells from a biological sample and for the in vitro stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GvH) inhibition, stimulation of graft versus tumour (GvT) and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. NKp30 antibodies are useful for identifying NKp30 natural ligands and allow assessment of the level of surface NKp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence NKp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the extracellular region of human NKp30 receptor

XX Sequence 120 AA;

Query Match 62.0%; Score 632; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 4.7e-54;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LWSVOPPEIRTLGSSAPLPCSFNASQGRLAIGSVTFRDEVVPGKEVRNGTPEFRGLA 78
DB 1 LWSVOPPEIRTLGSSAPLPCSFNASQGRLAIGSVTFRDEVVPGKEVRNGTPEFRGLA 60

QY 79 PLASSRFLHDHQAELHIRDVRGHDASTYVCRVEVLGLGVGTGTLVVEKEHPOLGAGT 138
DB 61 PLASSRFLHDHQAELHIRDVRGHDASTYVCRVEVLGLGVGTGTLVVEKEHPOLGAGT 120

RESULT 8
AAE02773
ID AAE02773 standard; peptide; 33 AA.

XX AAE02773;
XX AC AAE02773;
XX DT 06-AUG-2001 (first entry)
XX DE Human NKp30 receptor intracellular region sequence.
XX DE Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;
XX KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
XX KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
XX KW therapy; intracellular region.
XX OS Homo sapiens.
XX PN WO200136630-A2.
XX PD 25-MAY-2001.
XX PF 15-NOV-2000; 2000WO-EP011697.
XX PR 15-NOV-1999; 99CA-02288307.
XX PR 15-NOV-1999; 99US-00440514.
XX XX (INNA-) INNATE PHARMA SAS.
XX PA (UYGE-) UNIV GENOVA.
XX PI Moretta A, Bottino C, Biassoni R;
XX XX WPI; 2001-329221/34.
XX DR Novel compound, useful for detection and/or quantifying the presence of
XX PT NK cells, comprises the amino acid sequences of the NKp30 molecule.
XX XX Claim 1; Fig 7B; 83pp; English.
XX The invention relates to human NKp30 receptor and its corresponding cDNA

CC molecule which is involved in natural cytotoxicity mediated by natural
 CC killer (NK) cells and antibodies that identify the same. Nkp30 receptor
 CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively
 CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are
 CC useful for detecting and/or quantifying the presence of NK cells in a
 CC biological sample. The invention also provides kits for detecting and/or
 CC quantifying the presence of NK cells, for the selective removal of NK
 CC cells from a biological sample, for the positive and selective
 CC purification of NK cells from a biological sample and for the in vitro
 CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GvH) inhibition, stimulation of graft
 CC versus tumour (GvT) and especially graft versus leukaemia (GvL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. Nkp30 antibodies are useful for
 CC identifying Nkp30 natural ligands and allow assessment of the level of
 CC surface Nkp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence Nkp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is the intracellular region of human Nkp30 receptor
 XX
 SQ Sequence 33 AA;

Query Match 20.0%; Score 204; DB 4; Length 33;
 Best Local Similarity 100.0%; Pred. No. 9.7e-13;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 STVYQKCHCHGTHCHSSDGRGVIPRCP 190
 DB 1 STVYQKCHCHGTHCHSSDGRGVIPRCP 33

RESULT 9
 AAE20273
 ID AAE20273 standard; protein; 246 AA.
 AC AAE20273;
 XX
 XX 18-JUN-2002 (first entry)
 DT
 DE Human lung specific gene (LSG) protein #11.
 XX
 XX Human; lung specific gene; LSG; lung embryonic development; cytostatic;
 KW lung cancer; vaccine; gene therapy; non-cancerous lung disease; tumour.
 KW
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 PH Region 189..204
 FT /note= "Antigenic epitope"
 FT
 XX WO200208278-A2.
 PN
 XX 31-JAN-2002.
 PD
 XX 20-JUL-2001; 2001WO-US022949.
 PF
 XX 21-JUL-2000; 2000US-0219834P.
 PR
 XX (DIAD-) DIADEXUS INC.
 PA
 XX Macina RA, Nair M, Chen S;
 PI
 XX WPI; 2002-268964/31.
 DR
 XX Novel lung specific gene useful for identifying, diagnosing, monitoring,
 PT staging, imaging and treating lung cancer and non-cancerous disease
 PT states in lung, for gene therapy, and for identifying lung tissue.
 PT
 PS Claim 2; Page 178-179; 197pp; English.
 XX
 CC The present invention relates to lung specific genes (LSG) and their

CC corresponding polypeptides. LSG is useful for identifying, diagnosing,
 CC monitoring, staging, imaging and treating lung cancer and non-cancerous
 CC disease states in lung, identifying lung tissue, monitoring and modifying
 CC lung embryonic development and differentiation, in gene therapy, as
 CC hybridisation probes, to detect LSG mRNA as a marker for lung cancer, as
 CC research reagents and materials for discovery of treatments and
 CC diagnostics to human disease, to detect complementary polynucleotides,
 CC and for chromosome identification. An antibody which binds LSG is useful
 CC to detect or image localisation of LSG in a patient for detecting or
 CC diagnosing a disease or condition, for preventing the onset and treatment
 CC of lung cancer, to isolate or to identify clones expressing LSG
 CC polypeptides, to purify LSG polypeptides, and to target tumours
 CC expressing LSG. The present sequence is human LSG protein
 XX
 SQ Sequence 246 AA;

Query Match 12.1%; Score 123.5; DB 5; Length 246;
 Best Local Similarity 29.2%; Pred. No. 0.001;
 Matches 54; Conservative 17; Mismatches 61; Indels 53; Gaps 10;

QY 1 MAMW-LILLIMVHFGSCALWV-SQPEITRTLSGSAFLPCSNASQGRLAIGSVTWFRD 58
 DB 1 MAMAPLLLTLLSLTGLSQFILTQPPSASALGASVTLTCSVSSDYKNL---EVDWFOQ 57

QY 59 EWPFGKEVR-----NGTPEFRGLAP-----LASSRFLHDHQAELHIRDVRGHAS 104
 DB 58 R--PKGPRFVMRVGTGVVGFAGADIPDRFSVSGSLARF-----LTIRIEDEDS 108

QY 105 IYCVREVLGLGVGT-----NGTSLVV-----EKEHPQLGAGTVLL 142
 DB 109 DYHCGTD---LGSSTFSVSWVFGGTLTLVLSQPKAAPSVTLPFPSPSEELQANKATLVCL 165

QY 143 RAGFY 147
 DB 166 ISDFY 170

RESULT 10
 ABU08020
 ID ABU08020 standard; protein; 232 AA.
 AC ABU08020;
 XX
 XX 10-MAY-2003 (first entry)
 DT
 DE Monoclonal rabies virus antibody light chain, clone JB.1.
 XX
 KW Human; antibody; constant region; monoclonal antibody 57; Mab 57;
 KW variable region; Rabies; neurological disease; infection;
 KW central nervous system; rabies virus; Lyssavirus; Rhabdoviridae;
 KW pathogen; vaccine; virucide; light chain.
 KW
 XX Homo sapiens.
 OS
 XX WO2003016501-A2.
 PN
 XX 27-FEB-2003.
 PD
 XX 21-AUG-2002; 2002WO-US026584.
 PF
 XX 21-AUG-2001; 2001US-0314023P.
 PR
 XX (UYJE-) UNIV JEFFERSON THOMAS.
 PA
 XX Hooper DC, Dietzschold B;
 PI
 XX WPI; 2003-278566/27.
 DR N-PSDB; ABX12862.
 XX
 PT New recombinant antibody comprising a constant region of Mab 57 linked to
 PT a non-Mab 57 variable region, useful for treating an individual exposed
 PT to a pathogen, e.g. rabies infection.
 XX


```

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX PA Linsley PS, Ledbetter JA, Damle NK, Brady W, Wallace PM;
XX PI
XX XX WPI; 1994-272160/34.
XX DR N-PSDB; AAQ70451.
XX DR
XX PT Use of a CTLA4-binding molecule and an IL4-binding molecule - for
XX PT regulating an immune response involving B7-positive lymphocytes, partic.
XX PT in transplant rejection.
XX
XX Example 3; Fig 3; 96pp; English.
XX
XX This sequence shows the human CTLA4 receptor. The cDNA encoding it was
XX CC assembled from 2 PCR fragments amplified from mRNA from H38 leukaemia
XX CC cells. The CTLA4 receptor is express using the the oncostatin M signal
XX CC peptide fused to the N terminus of CTLA4. The methods of the invention
XX CC can be used for regulating a B cell response resulting in the inhibition
XX CC of antibody prodn., regulating a T cell response resulting in the
XX CC inhibition of cell mediated immunity or for the inhibition of lymphocyte
XX CC proliferation (Claimed)
XX
XX Sequence 212 AA;
XX
XX Query Match 11.6%; Score 118.5; DB 2; Length 212;
XX Best Local Similarity 28.1%; Pred. No. 0.0026;
XX Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;
XX
XX QY 4 MLLLLILIMVHP--GSCALWVSQPPEIRLTGSSAFLPCSFNASQGRLAIGSVTWFDRDEW 61
XX Db 10 LLSLLVLLLFPSMASMAHVAQPAVLASSRGIAFVCEY-ASPGKATEVRVTILRQADS 68
XX
XX QY 62 PCGEVRENTPEFRGELAPLASSRLHP-----HQAELHTRDVRGHDSIYCVREV 112
XX Db 69 QVTEVCAATYMGNELT-----FLDSICTGTSSGNQNLITQGLRAMDTGLYICKVEL 122
XX
XX QY 113 L---GLGVGTGNTGLRVVEKEHPQLGAGTVLLLRA----GFYAVSFLSVAV 156
XX Db 123 MYPPPPYLIGNGTGIYVIDPEPCPSDFLLWILAAVSSGLPFYSFLTAV 173
XX
XX RESULT 13
XX AAR77642
XX ID AAR77642 standard; protein; 212 AA.
XX AC
XX AC AAR77642;
XX XX
XX DT 11-JUN-1996 (first entry)
XX
XX DE Full length CTLA4 and oncostatin M signal peptide.
XX
XX KW primer; PCR; polymerase chain reaction; CTLA4; CD28; B7 antigen;
XX KW fusion protein; regulation; immune response; lymphocyte; oncostatin M;
XX KW soluble; chimeric; inhibit; transplant rejection;
XX KW graft versus host disease.
XX XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..25
XX FT /label= oncostatin_M_signal_peptide
XX FT Protein 26
XX FT /label= CTLA4
XX FT Modified-site 134..136
XX FT /note= "glycosylation site"
XX
XX XX AU9516458-A.
XX DN
XX XX 26-OCT-1995.
XX PD
XX XX 13-APR-1995; 95AU-00016458.
XX PF
XX XX

```


PF 04-OCT-1996; 96US-00725776.
 XX 27-JUN-1991; 91US-00723617.
 PR 22-JAN-1993; 93US-00008898.
 PR 15-APR-1994; 94US-00228208.
 PR 18-JAN-1995; 95US-00375390.
 PR 03-JUN-1995; 95US-00465078.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Ledbetter JA, Damle NK, Brady W, Kiener PA, Linsley PS;
 PI WPI; 1999-600811/51.
 DR N-PSDB; AAZ29989.
 XX Regulating CTLA4 positive T cell interactions.
 PT Example 3; Fig 3; 75pp; English.
 XX The present sequence represents the human CTLA4 protein. The protein is
 CC used in the course of the invention. The specification describes a method
 CC for regulating CTLA4 receptor positive T cell interactions with B7
 CC receptor positive B cells. The method comprises contacting the CTLA4-
 CC positive T cells with monoclonal antibody fragments reactive with CTLA4.
 CC This inhibits (and therefore regulates) interactions between CTLA4-
 CC positive T cells and B7 positive B cells. The method may be used for
 CC regulating CTLA4 receptor positive T cell interactions with B7 receptor
 CC positive B cells. In this way the immune system of an individual can be
 CC manipulated (especially suppressed) for the treatment of autoimmune
 CC diseases (especially lupus erythematosus) and to prevent host-graft and
 CC transplant rejection
 XX Sequence 212 AA;
 SQ
 Query Match 11.6%; Score 118.5; DB 2; Length 212;
 Best Local Similarity 28.1%; Pred. No. 0.0026;
 Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;
 QY 4 MLLILIMVHP--GSCALWVSQPPEIRTLGSSAFPLCSFNASQGRLAIGSVTWFRDEVV 61
 Db 10 LLSVLALLPFSMASMAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADS 68
 QY 62 PKKEVRNGTPEFRGRAPLASSRFLHD-----HOAELHIRDVRGHDSIYVCREV 112
 Db 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLITQGLRAMDTGLYICKVEL 122
 QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLLR-----GFYAVSFLSVAV 156
 Db 123 MYPPPYLIGNGTQIVVIDPEPCPSDFLLWILAAVSSGLFFYSFLLTAV 173
 RESULT 15
 AAW81584
 ID AAW81584 standard; protein; 212 AA.
 AC AAW81584;
 XX
 DT 05-FEB-1999 (first entry)
 XX Human CTLA4 receptor polypeptide.
 XX CTLA4 receptor; CTLA4-Ig; fusion protein; B7 antigen; hinge; CH2; CH3;
 KW extracellular domain; human; immunoglobulin; T cell; immune system;
 KW autoimmune disease; cancer; viral infection.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..25
 FT Protein /note= "oncostatin M signal peptide"
 FT Protein 26..212
 FT Domain /note= "mature CTLA4 protein"
 FT Domain 26..150

/note= "extracellular domain used for constructing a
 CTLA4-Ig fusion protein"
 134..136
 /note= "Asn is N-glycosylated"
 Modified-site
 US844095-A.
 01-DEC-1998.
 18-JAN-1995; 95US-00375390.
 27-JUN-1991; 91US-00723617.
 22-JAN-1993; 93US-00008898.
 28-MAY-1993; 93US-00069693.
 15-APR-1994; 94US-00228208.
 (BRIM) BRISTOL-MYERS SQUIBB CO.
 Brady W, Linsley PS, Damle NK, Ledbetter JA;
 WPI; 1999-044666/04.
 N-PSDB; AAV69786.
 Fusion protein of CTLA4 and immunoglobulin fragment - for treating immune
 system disorders.
 Claim 2; Fig 3; 75pp; English.
 This represents a human CTLA4 receptor. The invention provides a CTLA4-Ig
 fusion protein that binds the B7 antigen and has a first amino acid
 sequence consisting of the extracellular domain of CTLA4 and a second
 amino acid sequence consisting of the hinge, CH2 and CH3 regions of a
 human immunoglobulin molecule. The fusion protein inhibits interaction of
 T cells with B7-positive cells and may be useful for treating immune
 system diseases, e.g. autoimmune diseases, cancer or viral infections
 Sequence 212 AA;
 Query Match 11.6%; Score 118.5; DB 2; Length 212;
 Best Local Similarity 28.1%; Pred. No. 0.0026;
 Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;
 QY 4 MLLILIMVHP--GSCALWVSQPPEIRTLGSSAFPLCSFNASQGRLAIGSVTWFRDEVV 61
 Db 10 LLSVLALLPFSMASMAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADS 68
 QY 62 PKKEVRNGTPEFRGRAPLASSRFLHD-----HOAELHIRDVRGHDSIYVCREV 112
 Db 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLITQGLRAMDTGLYICKVEL 122
 QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLLR-----GFYAVSFLSVAV 156
 Db 123 MYPPPYLIGNGTQIVVIDPEPCPSDFLLWILAAVSSGLFFYSFLLTAV 173
 Search completed: February 26, 2004, 12:10:03
 Job time : 121.131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 66.0212 Seconds
(without alignments)
607.670 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 MAWMLLLILIMVHPGSCALW.....GTHCHSSDGPGRGVIPEPRCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum March 1968
Listing first 45 summaries

Database : Published Applications AA:*

```

1: /csm2_6/pdata/2/pubaa/us07 PUBCOMB.pcp.*
2: /csm2_6/pdata/2/pubaa/FCT_NEW PUB.pcp.*
3: /csm2_6/pdata/2/pubaa/us06 NEW PUB.pcp.*
4: /csm2_6/pdata/2/pubaa/us06 PUBCOMB.pcp.*
5: /csm2_6/pdata/2/pubaa/us07 NEW PUB.pcp.*
6: /csm2_6/pdata/2/pubaa/FCTUS PUBCOMB.pcp.*
7: /csm2_6/pdata/2/pubaa/us08 NEW PUB.pcp.*
8: /csm2_6/pdata/2/pubaa/us09A PUBCOMB.pcp.*
9: /csm2_6/pdata/2/pubaa/us09B PUBCOMB.pcp.*
10: /csm2_6/pdata/2/pubaa/us09C PUBCOMB.pcp.*
11: /csm2_6/pdata/2/pubaa/us09 NEW PUB.pcp.*
12: /csm2_6/pdata/2/pubaa/us10A PUBCOMB.pcp.*
13: /csm2_6/pdata/2/pubaa/us10B PUBCOMB.pcp.*
14: /csm2_6/pdata/2/pubaa/us10C PUBCOMB.pcp.*
15: /csm2_6/pdata/2/pubaa/us10 NEW PUB.pcp.*
16: /csm2_6/pdata/2/pubaa/us10 PUBCOMB.pcp.*
17: /csm2_6/pdata/2/pubaa/us60 NEW PUB.pcp.*
18: /csm2_6/pdata/2/pubaa/us60 PUBCOMB.pcp.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	1020	100.0	190	13	US-10-036-444-2	Sequence 2, Appli
2	532	60.0	120	13	US-10-036-444-4	Sequence 4, Appli
3	504	28.0	33	13	US-10-036-444-6	Sequence 6, Appli
4	123.5	12.1	246	10	US-09-309-567B-49	Sequence 49, Appli
5	119.5	11.7	232	14	US-10-325-108A-12	Sequence 12, Appli
6	119.5	11.7	232	15	US-10-461-148-6	Sequence 6, Appli
7	118.5	11.6	139	14	US-10-312-495-6	Sequence 6, Appli
8	118.5	11.6	212	10	US-09-998-195A-17	Sequence 17, Appli
9	118.5	11.6	212	14	US-10-057-288-12	Sequence 12, Appli
10	118.5	11.6	212	14	US-10-455-514-2	Sequence 2, Appli
11	118.5	11.6	212	16	US-10-419-008-17	Sequence 17, Appli
12	111.5	10.9	226	13	US-10-038-107A-1	Sequence 1, Appli
13	111.5	10.9	226	15	US-10-371-069-32	Sequence 32, Appli
14	111.5	10.9	226	15	US-10-371-645-32	Sequence 32, Appli
15	111.5	10.9	226	15	US-10-371-260-32	Sequence 32, Appli

16	111.5	10.9	226	15	US-10-411-010-19	Sequence 19, Appl
17	110	10.8	223	9	US-09-989-545-21	Sequence 21, Appl
18	110	10.8	223	10	US-09-928-267-18	Sequence 18, Appl
19	110	10.8	223	10	US-09-928-267-21	Sequence 21, Appl
20	110	10.8	223	14	US-09-928-267-18	Sequence 8, Appl
21	110	10.8	223	14	US-10-225-519-8	Sequence 8, Appl
22	110.5	10.8	223	14	US-10-207-655-101	Sequence 101, App
23	108.5	10.6	223	9	US-09-989-545-20	Sequence 20, Appl
24	108.5	10.6	223	14	US-10-211-207-5	Sequence 5, Appl
25	108.5	10.6	223	14	US-10-077-106-5	Sequence 5, Appl
26	108.5	10.6	223	14	US-09-898-195A-11	Sequence 11, Appl
27	108.5	10.6	383	14	US-10-057-288-6	Sequence 6, Appl
28	108.5	10.6	383	14	US-10-155-514-10	Sequence 10, Appl
29	108	10.6	223	13	US-10-419-008-11	Sequence 11, Appl
30	108	10.6	223	13	US-10-107-828-26	Sequence 26, Appl
31	108	10.6	223	13	US-10-107-907-26	Sequence 26, Appl
32	108	10.6	223	13	US-10-107-868-26	Sequence 26, Appl
33	108	10.6	223	14	US-10-301-058-26	Sequence 26, Appl
34	108	10.6	260	10	US-09-928-267-13	Sequence 13, Appl
35	108	10.6	260	14	US-09-928-267-14	Sequence 14, Appl
36	107.5	10.5	261	10	US-10-225-519-4	Sequence 4, Appl
37	107.5	10.5	261	10	US-09-928-267-9	Sequence 9, Appl
38	107.5	10.5	261	14	US-09-928-267-10	Sequence 10, Appl
39	107	10.5	223	10	US-10-225-519-2	Sequence 2, Appl
40	107	10.5	223	14	US-09-835-297-2	Sequence 2, Appl
41	107	10.5	223	14	US-10-211-207-3	Sequence 3, Appl
42	107	10.5	223	14	US-10-077-106-3	Sequence 3, Appl
43	107	10.5	382	15	US-10-336-384-7	Sequence 7, Appl
44	106.5	10.4	151	14	US-10-207-655-307	Sequence 307, App
45	106.5	10.4	223	10	US-09-928-267-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

```

US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "NO. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-2

```

```
Query Match      100.0%; Score 1020; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 2e-94;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	MAWMLLLILIMVHPGSCALWVQSPPEIITLGGSAFLPCSFNAGQRLAIGSVTWF	60
Db	1	MAWMLLLILIMVHPGSCALWVQSPPEIITLGGSAFLPCSFNAGQRLAIGSVTWF	60
Qy	61	VPGKEVRNGTPEFRGLAPLASSRFLPHQAEHLIRDVRGHDASIYVCRVEVLGLGVGTG	121
Db	61	VPGKEVRNGTPEFRGLAPLASSRFLPHQAEHLIRDVRGHDASIYVCRVEVLGLGVGTG	121
Qy	121	NGETRLWEKEHPOLGAGTVLLLRAGFYAVSFLIVAGSVTVYOGKCHCMGTGCHSSDGP	181

Db 121 NGRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAGSVITYYQKCHCHGTHCHSSDGP 180
QY 181 RGVIPERCP 190
Db 181 RGVIPERCP 190

RESULT 2
US-10-036-444-4
; Sequence 4, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-PR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-4

Query Match 62.0%; Score 632; DB 13; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 LWSQPEITRLTSGSAFLPCSFNASQGLAIGSVTWFRDEVVPGKEVNGTPEERGLA 78
Db 1 LWSQPEITRLTSGSAFLPCSFNASQGLAIGSVTWFRDEVVPGKEVNGTPEERGLA 60
QY 79 PLASSRFLDHQAEHLHIRDVRGHDAIYVCRVEVLGLVGTGNGRLVVEKEHPQLGAGT 138
Db 61 PLASSRFLDHQAEHLHIRDVRGHDAIYVCRVEVLGLVGTGNGRLVVEKEHPQLGAGT 120

RESULT 3
US-10-036-444-6
; Sequence 6, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-PR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-6

Query Match 20.0%; Score 204; DB 13; Length 33;
Best Local Similarity 100.0%; Pred. No. 2e-13;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 STVYQKCHCHGTHCHSSDGRGVIPERCP 190
Db 1 STVYQKCHCHGTHCHSSDGRGVIPERCP 33

RESULT 4
US-09-909-567B-49
; Sequence 49, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Selyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 49
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-49

Query Match 12.1%; Score 123.5; DB 10; Length 246;
Best Local Similarity 29.2%; Pred. No. 0.00031;
Matches 54; Conservative 17; Mismatches 61; Indels 53; Gaps 10;
QY 1 MAMM-LLILIMVHPGSCALMV-SQPEITRLTSGSAFLPCSFNASQGLAIGSVTWFRD 58
Db 1 MAMAPLLLTLSLTGSLQSPILTPPSASASLGASVLTCTSVSSDYKLN---EVDWFOQ 57
QY 59 EVVPGKEVR-----NGTPEERGLAP-----LASSRFLDHQAEHLHIRDVRGHDA 104
Db 58 R--PGKGRFVVRVGTGVVGRGADIPDRFSVSGSLNRF-----LTIRNIEEDES 108
QY 105 IYVCRVEVLGLVGT-----GNGTRLVY-----EKEHPQLGAGTVLL 142
Db 109 DYHCGTD--LGSSTSVSVWVFGGKTLVLSQPKAAPSVTLPFPSPSEELQANKATIVCL 165
QY 143 RAGFY 147
Db 166 ISDFY 170

RESULT 5
US-10-225-108A-12
; Sequence 12, Application US/10225108A
; Publication No. US20030157112A1
; GENERAL INFORMATION:
; APPLICANT: HOOPER, Craig
; APPLICANT: DIETZSCHOLD, Bernhard
; TITLE OF INVENTION: Recombinant Antibodies, and Compositions
; TITLE OF INVENTION: and Methods for Making Them
; FILE REFERENCE: 8321-110
; CURRENT APPLICATION NUMBER: US/10/225,108A
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 09/848,832
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/204,518
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/314,023
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 232
; TYPE: PRT

; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A099PCT000454-124

RESULT 9

US-10-057-288-12
; Sequence 12, Application US/10057288
; Publication No. US20030007968A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Christian P.
; APPLICANT: Pearson, Thomas C.
; APPLICANT: Waller, Edmund K.
; APPLICANT: Adams, Andrew B.

; TITLE OF INVENTION: METHODS OF INDUCING ORGAN TRANSPLANT TOLERANCE AND
; FILE REFERENCE: CORRECTING HEMOGLOBINOPATHIES
; CURRENT APPLICATION NUMBER: US/10/057,288
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/264,528
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/303,142
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 12
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-288-12

Query Match 11.6%; Score 118.5; DB 14; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWVSQPEIETLEGSAFLPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVALLFPMSMASMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 68

QY 62 PGKEVRNGTPEFRGRAPLASSRFLHD-----HQAELHIRDVRGHDASIVYCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA-----GFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

RESULT 10

US-10-155-514-2
; Sequence 2, Application US/10155514
; Publication No. US20030022836A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Christian P.
; APPLICANT: Pearson, Thomas C.
; APPLICANT: Adams, Andrew B.
; TITLE OF INVENTION: METHODS FOR PROTECTING ALLOGENEIC ISLET TRANSPLANT USING SOLUBLE
; FILE REFERENCE: MUTANT MOLECULES
; CURRENT APPLICATION NUMBER: US/10/155,514
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/293,402
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-155-514-2

Query Match 11.6%; Score 118.5; DB 14; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWVSQPEIETLEGSAFLPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVALLFPMSMASMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 68

QY 62 PGKEVRNGTPEFRGRAPLASSRFLHD-----HQAELHIRDVRGHDASIVYCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA-----GFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

RESULT 11

US-10-419-008-17
; Sequence 17, Application US/10419008
; Publication No. US20040022787A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; APPLICANT: Carr, Suzette
; APPLICANT: Hagerty, David
; APPLICANT: Peach, Robert J.
; APPLICANT: Becker, Jean-Claude
; TITLE OF INVENTION: METHODS FOR TREATING AN AUTOIMMUNE DISEASE USING A SOLUBLE CTLA4
; FILE REFERENCE: MOLECULE AND A DMARD OR NSAID
; CURRENT APPLICATION NUMBER: 30436.55US11
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 09/896,195
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,913
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/373,852
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/407,246
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-419-008-17

Query Match 11.6%; Score 118.5; DB 16; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWVSQPEIETLEGSAFLPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVALLFPMSMASMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 68

QY 62 PGKEVRNGTPEFRGRAPLASSRFLHD-----HQAELHIRDVRGHDASIVYCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA-----GFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

RESULT 12

US-10-038-107A-1
; Sequence 1, Application US/10038107A
; Publication No. US20020150573A1
; GENERAL INFORMATION:
; APPLICANT: Nussenzweig, Michel
; TITLE OF INVENTION: ANTI-1G ALPHA-BETA ANTIBODY FOR LYMPHOMA THERAPY
; FILE REFERENCE: 7529/OH405
; CURRENT APPLICATION NUMBER: US/10/038,107A
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/247,079
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

```
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-107A-1

Query Match      10.9%; Score 111.5; DB 13; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 LFLLSAVYLGPGCOALWMHKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWR--VLH 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLHIDVRGHDASIYVCRVEVLGLGVGTG 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 G-----NYTWPEFLGP-----GDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 CGTYLRVRQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLFLFR 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 13
US-10-371-069-32
; Sequence 32, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371,069
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-069-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 LFLLSAVYLGPGCOALWMHKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWR--VLH 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLHIDVRGHDASIYVCRVEVLGLGVGTG 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 G-----NYTWPEFLGP-----GDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 CGTYLRVRQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLFLFR 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 14
US-10-371-645-32
; Sequence 32, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-645-32

Query Match      10.9%; Score 111.5; DB 13; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 LFLLSAVYLGPGCOALWMHKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWR--VLH 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLHIDVRGHDASIYVCRVEVLGLGVGTG 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 G-----NYTWPEFLGP-----GDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 CGTYLRVRQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLFLFR 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
; Sequence 32, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-645-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 LFLLSAVYLGPGCOALWMHKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWR--VLH 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLHIDVRGHDASIYVCRVEVLGLGVGTG 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 G-----NYTWPEFLGP-----GDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 CGTYLRVRQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLFLFR 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 15
US-10-371-260-32
; Sequence 32, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-260-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY      4 MLLILLIMVHPGSCALWVSQ-PPEIRTLGSSAFLPCSFNASOGRLAIGSVTWERDEVVP 62
Db      18 LFLLSAVYLGCGQCALMWHKVPASLMSVLGEDAHFQCPHNSNN-----ANVTWNR-VLH 71

QY      63 GKEVRNGT--PEPRGRPLASSRFLHDHQAELHIEDVRGHDASIYVCREVVLGLGVGTG 120
Db      72 G----NYTWPPPEFLGP-----GDPNGTLIIQNANKSHGGIYVCRVQEGNESYQOS 118

QY      121 NGTRLVVEKEHPQ----LGGAGTV-LLLRAGFYAVSFLSVAGSVTVYQ 163
Db      119 CGYLRVRQPPRPFFLDMGEGTKNRIITAGIILLFCVAVPGTLLIFR 166
```

Search completed: February 26, 2004, 12:33:51
Job time : 68.0212 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 36.7905 Seconds
(without alignments)
266.616 Million cell updates/sec

Title: US-10-036-444-2
Perfect score: 1020
Sequence: 1 MAWMLLLILIMVHPSGSCALW.....GTHCHSDGPRVIEPRCP 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PCURS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	112	11.0	223	3	US-08-228-208A-17
2	111.5	10.9	226	4	US-09-311-784A-32
3	106.5	10.4	364	4	US-09-472-087-100
4	106	10.4	270	3	US-09-082-593-10
5	104.5	10.2	238	4	US-09-227-595-30
6	104.5	10.2	238	4	US-09-227-595-32
7	104	10.2	174	3	US-08-804-180C-4
8	103.5	10.1	187	1	US-08-067-684-14
9	103.5	10.1	187	1	US-08-008-898-14
10	103.5	10.1	187	2	US-08-459-818-14
11	103.5	10.1	187	2	US-08-889-666-14
12	103.5	10.1	187	2	US-08-465-078-14
13	103.5	10.1	187	2	US-08-725-776-14
14	103.5	10.1	187	2	US-08-488-062-14
15	103.5	10.1	187	3	US-08-228-208A-14
16	103.5	10.1	187	3	PCT-US95-06726-36
17	103.5	10.1	236	3	US-09-049-672A-7
18	103	10.1	267	1	US-08-416-336-2
19	102.5	10.0	223	4	US-09-303-040-10
20	101.5	10.0	374	4	US-09-327-595-26
21	101.5	10.0	374	4	US-09-327-595-28
22	101.5	10.0	377	4	US-09-327-595-24
23	98.5	9.7	526	4	US-08-910-174B-9
24	98.5	9.7	526	4	US-09-620-461-9
25	98	9.6	589	2	US-08-724-394A-1
26	97	9.5	131	1	US-08-305-683A-4
27	97	9.5	253	2	US-08-459-818-20

Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 21, Appl
Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-228-208A-17
; Sequence 17, Application US/08228208A
; Patent No. 6090914
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Wallace, Philip M.
; TITLE OF INVENTION: CTLA4/CD28ig HYBRID FUSION
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/228,208A
; FILING DATE: 15-APR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/008,898
; FILING DATE: 22-JAN-1993
; APPLICATION NUMBER: 07/723,617
; FILING DATE: 27-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-30US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-228-208A-17

Query Match 11.0%; Score 112; DB 3; Length 223;
Best Local Similarity 28.0%; Pred. No. 0.00046;
Matches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;
Qy 6 LLLILIMVHSGC-ALWISQPEIRTELGSSAFPLPCSFNASQGLAIGSVTWFRDEVPGK 64
Db 24 LLLFLLFIPVCKAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADSQVT 82
Qy 65 EVRNGTPEFRGLAPLASSRFLHD-----HQAELHIRDVRGHDAIYVCRVEVL-- 113
Db 83 EVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYP 136
Qy 114 -GLGVGTGNGTRLVVEKHPQLGAGTVLLRA-----GFYAVSFLSVAV 156
Db 137 PPTVGLGNGTQIYVIDPEPCDSDFLWILAAVSSGLFFYSFLLTAV 184

RESULT 2

US-09-311-784A-32
; Sequence 32; Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; TITLE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-09-311-784A-32

Query Match 10.9%; Score 111.5; DB 4; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.00053;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;
Qy 4 MLLILIMVHSGCALWVSQ-PPEIRTELGSSAFPLPCSFNASQGLAIGSVTWFRDEVVP 62
Db 18 LFLLSAVYLGPGCOALMWHKVPASLVSLGEDAHFQCPHNSNN-----ANVTWR--VLH 71
Qy 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
Db 72 G-----NVTWPEFLGP-----GEDPNTGTLIQNVKSHGGIYVCRVQEGNESYQOS 118
Qy 121 NGRVLVVEKHPQ-----LGAGTV-LLRAGFYAVSFLSVAGSVTVYIQ 163
Db 119 CGTVLRVRQPPRPFLDMGSGTKNRIITAGSIIILFCVAVPGTLLIFR 166

RESULT 3

US-09-472-087-100
; Sequence 100; Application US/09472087
; Patent No. 6682736
; GENERAL INFORMATION:
; APPLICANT: HANSON, DOUGLAS C.
; APPLICANT: NEVEU, MARK J.
; APPLICANT: MUELLER, EILEEN E.
; APPLICANT: HANKE, JEFFREY H.

; APPLICANT: GILMAN, STEVEN C.
; APPLICANT: DAVIS, C. GEOFFREY
; APPLICANT: CORVALAN, JOSE R.
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4
; FILE REFERENCE: ABX-PFI
; CURRENT APPLICATION NUMBER: US/09/472,087
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/113,647
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 100
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-472-087-100

Query Match 10.4%; Score 106.5; DB 4; Length 364;
Best Local Similarity 28.0%; Pred. No. 0.0035;
Matches 40; Conservative 20; Mismatches 62; Indels 21; Gaps 5;
Qy 4 MLLILIMVHP--GSCALWVSQPEIRTELGSSAFPLPCSFNASQGLAIGSVTWFRDEVV 61
Db 10 LLSLVIALLLFPFSMASMAHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADS 68
Qy 62 PKKEVRNGTPEFRGLAPLASSRFLHD-----HQAELHIRDVRGHDAIYVCRVEV 112
Db 69 QVTEVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122
Qy 113 L---GLGVGTGNGTRLVVEKEHP 132
Db 123 MYPPPYLIGNGTQIYVIDPEP 145

RESULT 4

US-09-082-593-10
; Sequence 10; Application US/09082593
; Patent No. 6180104
; GENERAL INFORMATION:
; APPLICANT: DAVIS, MARK M.
; APPLICANT: HEDRICK, STEPHEN M.
; TITLE OF INVENTION: T CELL RECEPTOR BETA SUBUNIT
; FILE REFERENCE: JX1193-195DIV2
; CURRENT APPLICATION NUMBER: US/09/082,593
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-082-593-10

Query Match 10.4%; Score 106; DB 3; Length 270;
Best Local Similarity 25.8%; Pred. No. 0.0026;
Matches 40; Conservative 22; Mismatches 45; Indels 48; Gaps 9;
Qy 3 MLLILIMVHSGCALWVS-----QPPEIRTELGSSAFPLPCSFNASQGLAIGSVTW 55
Db 12 WLL-----NWVNSQQVQSPESLIVPEGARTSLNCTFSDSASQY----FWW 55
Qy 56 PRDEVVPKEVR-----NGTPEFRGLAPLASSRFLHDQAELHIRDVRGHDAIYVC 108
Db 56 YRQH--SGKAPKALMSIFSNGEKE-EGRFTIHLNKASLH---FSLHIRDSPSDSALYLC 109
Qy 109 RVEVLGLGVGTGNGTRLVVEKEHPQLGAGTVLLIR 143
Db 110 AVTLVG-----GSGNKLI-----FGTGILLSVK 132

RESULT 5

US-09-227-595-30
; Sequence 30; Application US/09227595

Patent No. 6444792
; GENERAL INFORMATION:
; APPLICANT: Gray, Gary S. et al.
; TITLE OF INVENTION: CTLA4-Immunoglobulin Fusion Proteins and Uses Therefor
; TITLE OF INVENTION: Having Modified Effector Functions and Uses Therefor
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/595,590
; FILING DATE: February 2, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-227-595-30

Query Match 10.2%; Score 104.5; DB 4; Length 238;
Best Local Similarity 24.7%; Pred. No. 0.0032;
Matches 43; Conservative 24; Mismatches 70; Indels 37; Gaps 7;

Qy 1 MAWMLLILIM-----VHPSGSCALWVSQPPEIRLTGSSAFLPCSFNASQGRLAIGSVTW 55
Db 1 MGWSCIILFLVATATGVH---SAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTY 56

Qy 56 FRDEVPGKEVRNGTPEFRGLRLAPLASSRFLHD-----HQAELHIRDVRGHDASTY 106
Db 57 LRQADSQVTEVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLY 110

Qy 107 VCRVEVL---GLGVGTGNGTRLVV-----EKEHPQLGAGTVLLLRAGFY 147
Db 111 ICKVELMYPYPYILGIGNGAQITVAAPSVFIFFPPSDQLKSGTASVVCVLLNIFY 164

RESULT 6
US-09-227-595-32
; Sequence 32, Application US/09227595
; Patent No. 6444792
; GENERAL INFORMATION:
; APPLICANT: Gray, Gary S. et al.
; TITLE OF INVENTION: CTLA4-Immunoglobulin Fusion Proteins
; TITLE OF INVENTION: Having Modified Effector Functions and Uses Therefor
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/595,590
; FILING DATE: February 2, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-09-227-595-32

Query Match 10.2%; Score 104.5; DB 4; Length 238;
Best Local Similarity 24.7%; Pred. No. 0.0032;
Matches 43; Conservative 24; Mismatches 70; Indels 37; Gaps 7;

Qy 1 MAWMLLILIM-----VHPSGSCALWVSQPPEIRLTGSSAFLPCSFNASQGRLAIGSVTW 55
Db 1 MGWSCIILFLVATATGVH---SAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTY 56

Qy 56 FRDEVPGKEVRNGTPEFRGLRLAPLASSRFLHD-----HQAELHIRDVRGHDASTY 106
Db 57 LRQADSQVTEVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLY 110

Qy 107 VCRVEVL---GLGVGTGNGTRLVV-----EKEHPQLGAGTVLLLRAGFY 147
Db 111 ICKVELMYPYPYILGIGNGAQITVAAPSVFIFFPPSDQLKSGTASVVCVLLNIFY 164

RESULT 7
US-08-804-180C-4
; Sequence 4, Application US/08804180C
; Patent No. 6107056
; GENERAL INFORMATION:
; APPLICANT: Martin K. Oaks
; TITLE OF INVENTION: sCTLA-4 and Its Soluble Products
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas M. Wozny
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch Disk
; COMPUTER: IBM
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,180C
; FILING DATE: February 20, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas M. Wozny
; REGISTRATION NUMBER: 28,922

Query Match 10.1%; Score 103.5; DB 1; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77
DB 1 AHVQAQPAVVLAASSRGIAFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113
QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156
DB 114 YVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 148

RESULT 10
US-08-459-818-14
; Sequence 14, Application US/08459818
; Patent No. 5851795
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FastSeq 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,818
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436.35US02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-818-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77
DB 1 AHVQAQPAVVLAASSRGIAFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113
QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77
DB 1 AHVQAQPAVVLAASSRGIAFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113
QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156

DB 114 YVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 148

RESULT 11
US-08-889-666-14
; Sequence 14, Application US/0889666
; Patent No. 5885579
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; APPLICANT: Kiener, Peter A.
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Blvd., Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,666
; FILING DATE: 08-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375390
; FILING DATE: 18-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 30436-35US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-889-666-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77
DB 1 AHVQAQPAVVLAASSRGIAFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113
QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156
DB 114 YVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 148

RESULT 12
US-08-465-078-14
; Sequence 14, Application US/08465078
; Patent No. 5885796
; GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.
APPLICANT: Ledbetter, Jeffrey A.
APPLICANT: Damle, Nitin K.
APPLICANT: Brady, William
APPLICANT: Kiener, Peter A.
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,078

FILING DATE: 05-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375390

FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:

NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30436-35US01
TELEPHONE: 310-445-1140

TELEFAX: 310-445-9031

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-465-078-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSNASQGRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGRL 77
Db 1 AMHVAQPAVVLASSRGIA SFVCEY-ASPGKATEVRVTVLQRQADSQVTEVCAATYMGNEL 59

QY 78 APLASSRFLHD-----HQAEIHIRDVRGHDASIVYCRVEVL-----GLGVGTGNGTRL 125
Db 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMPPPYLGGNGTQI 113

QY 126 VVEKEHPOLGAGTVLLRA-----GFYAVSFLSVAV 156
Db 114 YVIDPEPCDSDFLLWILAAVSSGLFFYSFLLTAV 148

RESULT 13

US-08-725-776-14
Sequence 14, Application US/08725776
Patent No. 5968510

GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.
APPLICANT: Ledbetter, Jeffrey A.
APPLICANT: Damle, Nitin K.
APPLICANT: Brady, William
APPLICANT: Kiener, Peter A.

TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Blvd., Suite 400

CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,776

FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375390

FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:

NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30436-35US01
TELEPHONE: 310-445-1140

TELEFAX: 310-445-9031

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-725-776-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSNASQGRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGRL 77
Db 1 AMHVAQPAVVLASSRGIA SFVCEY-ASPGKATEVRVTVLQRQADSQVTEVCAATYMGNEL 59

QY 78 APLASSRFLHD-----HQAEIHIRDVRGHDASIVYCRVEVL-----GLGVGTGNGTRL 125
Db 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMPPPYLGGNGTQI 113

QY 126 VVEKEHPOLGAGTVLLRA-----GFYAVSFLSVAV 156
Db 114 YVIDPEPCDSDFLLWILAAVSSGLFFYSFLLTAV 148

RESULT 14

US-08-488-062-14
Sequence 14, Application US/08488062
Patent No. 5977318

GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.
APPLICANT: Ledbetter, Jeffrey A.
APPLICANT: Damle, Nitin K.
APPLICANT: Brady, William
APPLICANT: Kiener, Peter A.

TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Blvd., Suite 400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90025

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,062
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375390
FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30436-35US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310-445-1140
TELEFAX: 310-445-9031
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-062-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;
QY 18 ALWVSQPEIRTELGSSAFPCSFNASQRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGL 77
Db 1 AMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVTVLRQADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEIHIDVRGHDASIYVCRVEVL---GLGVGTGNGTRL 125
Db 60 T-----FLDDSICTGTSGNQVNLTIGLRAMDTGLYICKVELMYPYPYVLGNGTQI 113
QY 126 VVEKEHPQAGTVLLRA-----GFYAVSFLSVAV 156
Db 114 YVIDPEPCDSDFLWILAAVSSGLFFYSFLLTAV 148

RESULT 15

US-08-228-208A-14
Sequence 14, Application US/08228208A
Patent No. 6090914
GENERAL INFORMATION:
APPLICANT: Linsley, Peter S.
APPLICANT: Ledbetter, Jeffrey A.
APPLICANT: Damle, Nitin K.
APPLICANT: Brady, William
APPLICANT: Wallace, Philip M.
TITLE OF INVENTION: CTLA4/CD28lg HYBRID FUSION
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merchant & Gould
STREET: 11150 Santa Monica Boulevard, Suite 400
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/228,208A
FILING DATE: 15-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/008,898
FILING DATE: 22-JAN-1993
APPLICATION NUMBER: 07/723,617
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:

NAME: Adriano, Sarah B.
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30436-30US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 310 445-1140
TELEFAX: 310 445-9031
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-228-208A-14

Query Match 10.1%; Score 103.5; DB 3; Length 187;
Best Local Similarity 27.7%; Pred. No. 0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;
QY 18 ALWVSQPEIRTELGSSAFPCSFNASQRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGL 77
Db 1 AMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVTVLRQADSQVTEVCAATYMMGNEL 59
QY 78 APLASSRFLHD-----HQAEIHIDVRGHDASIYVCRVEVL---GLGVGTGNGTRL 125
Db 60 T-----FLDDSICTGTSGNQVNLTIGLRAMDTGLYICKVELMYPYPYVLGNGTQI 113
QY 126 VVEKEHPQAGTVLLRA-----GFYAVSFLSVAV 156
Db 114 YVIDPEPCDSDFLWILAAVSSGLFFYSFLLTAV 148

Search completed: February 26, 2004, 12:16:14
Job time : 37.7905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 51.2467 Seconds
(without alignments)
738.822 Million cell updates/sec

Title: US-10-036-444-4

Perfect score: 632

Sequence: 1 LNWSPPEIRTEGSSAFPLP.....TGNTRLVVEKHPOLGAGT 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp archea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp rvirus:*
- 16: sp bacteriap:*
- 17: sp archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632	100.0	177	014930	014930 homo sapien
2	632	100.0	190	014932	014932 homo sapien
3	632	100.0	201	014931	014931 homo sapien
4	599	94.8	176	095JB8	Q95JB8 macaca fasc
5	599	94.8	180	Q8MJ02	Q8MJ02 macaca fasc
6	476.5	75.4	152	095668	Q95668 homo sapien
7	476.5	75.4	165	095669	Q95669 homo sapien
8	476.5	75.4	176	095667	Q95667 homo sapien
9	466	73.7	115	Q8MJ00	Q8MJ00 macaca mula
10	454.5	71.9	151	Q8MJ01	Q8MJ01 macaca mula
11	428	67.7	192	11 Q8CFD9	Q8CFD9 rattus norv
12	424	67.1	192	11 Q8QW88	Q8QW88 rattus norv
13	420	66.5	192	11 Q8CG11	Q8CG11 rattus norv
14	108.5	17.2	449	5 Q9NKA5	Q9NKA5 drosophila
15	105.5	16.7	235	11 Q99M11	Q99M11 mus musculu
16	103	16.3	329	4 Q8N225	Q8N225 homo sapien

17	103	16.3	1340	4	Q8NDA2	Q8nda2 homo sapien
18	101.5	16.1	221	6	Q28090	Q28090 bos taurus
19	100.5	15.9	221	6	Q97631	Q97631 ovis aries
20	96.5	15.3	350	5	Q9VFU7	Q9vfuf drosophila
21	94	14.9	526	4	Q9H458	Q9h458 homo sapien
22	93.5	14.8	274	11	Q80YU5	Q80yu5 mus musculu
23	92.5	14.6	223	11	Q7TWX1	Q7twx1 mus musculu
24	92	14.6	174	11	Q921A7	Q921a7 rattus norv
25	91.5	14.5	223	6	Q9TT02	Q9tt02 canis famil
26	91.5	14.5	223	6	Q9GKP2	Q9gkp2 canis famil
27	91.5	14.5	223	11	Q62859	Q62859 rattus norv
28	90.5	14.3	272	11	Q70356	Q70356 mus musculu
29	90	14.2	524	11	Q921K7	Q921k7 mus musculu
30	90	14.2	700	11	Q7TSU7	Q7tsu7 mus musculu
31	89.5	14.2	137	4	Q95653	Q95653 homo sapien
32	89	14.1	108	13	Q8UJ25	Q8uj25 orectolobus
33	88.5	14.0	223	6	Q9XTA1	Q9xta1 felis silve
34	88.5	14.0	223	6	Q9XSY7	Q9xsy7 felis silve
35	88.5	14.0	244	6	Q7YRD9	Q7yrd9 macaca fasc
36	88.5	14.0	247	6	Q9BGS7	Q9bgs7 macaca fasc
37	88	13.9	178	11	Q9D1U4	Q9diu4 mus musculu
38	88	13.9	179	11	Q921X1	Q921x1 mus musculu
39	88	13.9	523	11	Q8K2H7	Q8k2h7 mus musculu
40	88	13.9	946	13	Q07153	Q07153 torpedo cal
41	87.5	13.8	141	4	Q9NU66	Q9nu66 homo sapien
42	87.5	13.8	160	4	Q8TDA6	Q8tda6 homo sapien
43	87.5	13.8	176	4	Q9NU65	Q9nu65 homo sapien
44	87.5	13.8	183	4	Q9NU64	Q9nu64 homo sapien
45	87.5	13.8	224	4	Q96KV1	Q96kv1 homo sapien

ALIGNMENTS

RESULT 1

014930 PRELIMINARY; PRT; 177 AA.

AC 014930

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE 1C7 precursor (1C7 protein).

GN 1C7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;

RX MEDLINE=96422187; PubMed=8824804;

RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weisman S.M.;

RT "Genes in a 220-KB region spanning the TNF cluster in human MHC.";

RL Genomics 31:215-222(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;

RA Nalabolu S.R., Raghunathan A., Weisman S.M.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=93272029; PubMed=8499947;

RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,

RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;

RT "Dense Alu clustering and a potential new member of the NF kappa B

family within a 90 kilobase HLA class III segment.";

RL Nat. Genet. 3:137-145(1993).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=96215741; PubMed=8629302;

RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,

RA Wallace A.F., Russell M.E.;

RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage


```

RESULT 3
O14931 PRELIMINARY; PRT; 201 AA.
AC O14931;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7 precursor (NCR3 protein).
GN 1C7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96422187; PubMed=8824804;
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";
RL Genomics 31:215-222(1996).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
[3]
RN [3]
RP SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA Laskey S., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
RT region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment.";
RL Nat. Genet. 3:137-145(1993).
[5]
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8629302;
RA Urans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392(1996).
[6]
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=9606565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RT TNF region.";
RL Immunogenetics 42:315-322(1995).
[7]
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208881; PubMed=7916655;
RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.E.;
RT "Lymphotoxin beta, a novel member of the TNF family that forms a
RT heteromeric complex with lymphotoxin on the cell surface.";
RL Cell 72:847-856(1993).
[8]
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=86016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotoxin and tumor necrosis factor genes: structure,
RT homology and chromosomal localization.";
RL Nucleic Acids Res. 13:6361-6373(1985).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=91086846; PubMed=1670639;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Riethmuller G., Weiss E.H.;
RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
RT NcoI polymorphism in the first intron of the human TNF-beta gene
RT correlates with a variant amino acid in position 26 and a reduced
RT level of TNF-beta production.";
RL J. Exp. Med. 173:209-219(1991).
[10]
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671667;
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RT "Haplotypic polymorphisms of the TNFB gene.";
RL Immunogenetics 33:50-53(1991).
[11]
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94362679; PubMed=8081366;
RA Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
RT histocompatibility complex that encodes a potential new member of the
RT I kappa B family of proteins.";
RL Hum. Mol. Genet. 3:793-795(1994).
[12]
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324911; PubMed=7601445;
RA Pealman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative
RT nuclear RNA helicase of the DEAD family.";
RL Genomics 26:210-218(1995).
[13]
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RT "Alternative splicing of the LST-1 gene located in the major
RT histocompatibility complex on human chromosome 6.";
RL DNA Seq. 8:155-160(1997).
[14]
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
[15]
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;
RA Shlina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
RT Ikb1 and MICA genes at the centromeric end of the HLA class I
RT region.";
RL Genomics 47:372-382(1998).
[16]
RN [16]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22389257; PubMed=12477932;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusica K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.I., Sheetz T.E.,
RA Brownstein N.A., Udell T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

```

RA Villalon D.K., Muzny D.M., Sodergren E.J., Liu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [17]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031137; AAB86579.1; -;
DR EMBL; AF129756; AAD18088.1; -;
DR EMBL; Y14768; CAA75063.1; -;
DR EMBL; BC052582; AAB52582.1; -;
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW SIGNAL.
FT CHAIN 1 61 POTENTIAL.
FT CHAIN 62 201 1C7.
FT SIGNAL 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;
SQ SEQUENCE 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;
Query Match 100.0%; Score 632; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.6e-60;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LWSQPPPEIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 79 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138
RESULT 4
Q95JB8 PRELIMINARY; PRT; 176 AA.
AC Q95JB8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nkp30 (Nkp30v1).
GN NCR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541; 9544;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis; TISSUE=Lymphoid;
RA Rizzi M., Blassoni R.;
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in
RT Macaca fascicularis lymphoid cells";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification
RT of Nkp46SD and Nkp30S.";

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278399; CAC41081.1; -;
DR EMBL; AY035215; AAK63117.1; -;
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;
Query Match 94.8%; Score 599; DB 6; Length 176;
Best Local Similarity 95.0%; Pred. No. 8.3e-57;
Matches 114; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 LWSQPPPEIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 79 PLSSSRFLRDQAEHLHIRDVRGHDAGIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138
RESULT 5
Q9MJ02 PRELIMINARY; PRT; 180 AA.
AC Q9MJ02;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nkp30.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification
RT of Nkp46SD and Nkp30S.";

NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93272029; PubMed=8499947;
 RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFKappaB
 RT family within a 90 kilobase HLA class III segment."; Nat.
 RL Nat. Genet. 3:137-145(1993).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96215741; PubMed=86239032;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcenci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts."; Transplantation
 RL 61:1387-1392(1996).
 [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96008565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region."; Immunogenetics 42:315-322(1995).
 [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface."; Cell
 RL 72:847-856(1993).
 [5]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization."; Nucleic Acids Res. 13:6361-6373(1985).
 [6]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
 RT Level of TNF-beta Production."; J. Exp. Med. 173:209-219(1991).
 [7]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=1671667; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RA "Haplotypic polymorphisms of the TNFB gene."; Immunogenetics 33:50-53(1991).
 [8]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins."; Hum. Mol. Genet. 3:793-799(1994).
 [9]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,
 RA van Zeven A., Coppieers W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
 RT Nuclear RNA Helicase of the D-E-A-D Family.";

Genomics 26:210-218(1995).
 [10]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6."; DNA Seq. 8:155-160(1997).
 [11]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing."; Genomics 45:591-600(1997).
 [12]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98149985; PubMed=9480751;
 RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IkBL and MICA genes at the centromeric end of the HLA class I
 RT region."; Genomics 47:372-382(1998).
 [13]
 RN EMBL; Y14768; CAA75067.1; -;
 DR GO; GO:0003793; P:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG Like; 1.
 SQ SEQUENCE 152 AA; 16393 MW; 42718746451F9ADC CRC64;
 Query Match 75.4%; Score 476.5; DB 4; Length 152;
 Best Local Similarity 79.2%; Pred. No. 1.3e-43;
 Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
 QY 1 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
 DB 19 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKE----- 65
 QY 61 PLASRFLLHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120
 DB 66 -----AELHIRDVRGHDAIYVCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 113
 RESULT 7
 O95669
 ID O95669 PRELIMINARY; PRT; 165 AA.
 AC O95669;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7F.
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFKappaB
 RT family within a 90 kilobase HLA class III segment."; Nat. Genet. 3:137-145(1993).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFKappaB
 RT family within a 90 kilobase HLA class III segment."; Nat. Genet. 3:137-145(1993).
 [2]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quiet W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392(1996).
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=96006565; PubMed=7590964;
RX Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RT TNF region.";
RL Immunogenetics 42:315-322(1995).
RN [4]
RN SEQUENCE FROM N.A.
RP MEDLINE=9320881; PubMed=7916655;
RX Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
RT "Lymphotoxin-beta: A new member of the TNF family that forms a
RT heteromeric complex with lymphotoxin on the cell surface.";
RL Cell 72:847-856(1993).
RN [5]
RN SEQUENCE FROM N.A.
RP MEDLINE=86016093; PubMed=2995927;
RX Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotoxin and tumor necrosis factor genes: structure,
RT homology and chromosomal localization.";
RL Nucleic Acids Res. 13:6361-6373(1985).
RN [6]
RN SEQUENCE FROM N.A.
RP MEDLINE=91086846; PubMed=1670638;
RX Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Rethmuller G., Weiss E.H.;
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
RT Level of TNF-beta Production.";
RL J. Exp. Med. 173:209-219(1991).
RN [7]
RN SEQUENCE FROM N.A.
RP MEDLINE=91139175; PubMed=1671667;
RX Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RA "Haplotypic polymorphisms of the TNFS gene.";
RT Immunogenetics 33:50-53(1991).
RN [8]
RN SEQUENCE FROM N.A.
RP MEDLINE=94362679; PubMed=8081366;
RX Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
RT histocompatibility complex that encodes a potential new member of the
RT I kappa B family of proteins.";
RL Hum. Mol. Genet. 3:793-799(1994).
RN [9]
RN SEQUENCE FROM N.A.
RP MEDLINE=95324911; PubMed=7601445;
RX Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
RT Nuclear RNA Helicase of the D-E-A-D Family.";
RL Genomics 26:210-218(1995).
RN [10]
RN SEQUENCE FROM N.A.
RP MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RT "Alternative splicing of the LST-1 gene located in the major
RT histocompatibility complex on human chromosome 6.";
RL DNA Seq. 8:155-160(1997).
RN [11]
RN SEQUENCE FROM N.A.
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
RN [12]
RN SEQUENCE FROM N.A.
RP MEDLINE=98149985; PubMed=9480751;
RX Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Teguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura I.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
RT IKBL and MICA genes at the centromeric end of the HLA class I
RT region.";
RL Genomics 47:372-382(1998).
RX EMBL; Y14768; CAA75068.1;
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;
SQ
Query Match 75.4%; Score 476.5; DB 4; Length 165;
Best Local Similarity 79.2%; Pred. No. 1.4e-43;
Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
QY 1 LWSQPPPIRTLEGSSAFPLPCSFNASQGLAIGSVTWFRDEVPGKEVNGTPEFRGELA 60
Db 19 LWSQPPPIRTLEGSSAFPLPCSFNASQGLAIGSVTWFRDEVPGKE----- 65
QY 61 PLASSRFLHDQAEHLHNDVRGHDASIVYCRVEVLGLVGVTGNGTRLVVKEHPQLGAGT 120
Db 66 -----AELHNDVRGHDASIVYCRVEVLGLVGVTGNGTRLVVKEHPQLGAGT 113
RESULT 8
O95667 PRELIMINARY; PRT; 176 AA.
AC O95667
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7e.
GN 1C7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the Nfkapab
RT family within a 90 kilobase HLA class III segment.";
RL Nat. Genet. 3:137-145(1993).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quiet W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392(1996).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;

```

RA Holzinger I., de Baey A., Messer G., Kick G., Zwierniza H.,
RA Weiss E.H.;
RA "Cloning and genomic characterization of LST1: a new gene in the human
RA TNF region.",
RA Immunogenetics 42:315-322(1995).
RA [4]
RA SEQUENCE FROM N.A.
RA MEDLINE=93208881; PubMed=7916655;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.,
RA "Lymphotoxin-beta: A new member of the TNF family that forms a
RA heteromeric complex with lymphotoxin on the cell surface.",
RA Cell 72:847-856(1993).
RA [5]
RA SEQUENCE FROM N.A.
RA MEDLINE=86016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.,
RA "Human lymphotoxin and tumor necrosis factor genes: structure,
RA homology and chromosomal localization.",
RA Nucleic Acids Res. 13:6361-6373(1985).
RA [6]
RA SEQUENCE FROM N.A.
RA MEDLINE=91086846; PubMed=1670638;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Riethmüller G., Weiss E.H.,
RA "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
RA Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
RA Correlates with a Variant Amino Acid in Position 26 and a Reduced
RA Level of TNF-beta Production.",
RA J. Exp. Med. 173:209-219(1991).
RA [7]
RA SEQUENCE FROM N.A.
RA MEDLINE=91139175; PubMed=1671667;
RA Abraham I.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RA "Haplotypic polymorphisms of the TNFB gene.",
RA Immunogenetics 33:50-53(1991).
RA [8]
RA SEQUENCE FROM N.A.
RA MEDLINE=94362679; PubMed=8081366;
RA Albertella M.R., Campbell D.R.;
RA "Characterization of a novel gene in the human major
RA histocompatibility complex that encodes a potential new member of the
RA I kappa B family of proteins.",
RA Hum. Mol. Genet. 3:793-799(1994).
RA [9]
RA SEQUENCE FROM N.A.
RA MEDLINE=95324911; PubMed=7601445;
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RA "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
RA Nuclear RNA Helicase of the D-E-A-D Family.",
RA Genomics 26:210-218(1995).
RA [10]
RA SEQUENCE FROM N.A.
RA MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RA "Alternative splicing of the LST-1 gene located in the major
RA histocompatibility complex on human chromosome 6.",
RA DNA Seq. 8:155-160(1997).
RA [11]
RA SEQUENCE FROM N.A.
RA MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RA "Complex expression pattern of the TNF region gene LST1 through
RA differential regulation, initiation, and alternative splicing.",
RA Genomics 45:591-600(1997).
RA [12]
RA SEQUENCE FROM N.A.
RA MEDLINE=98149985; PubMed=9480751;
RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,

RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
RA Kimura M., Inoko H.;
RA "Nucleotide sequencing analysis of the 146-kilobase segment around the
RA IKB1 and MICA genes at the centromeric end of the HLA class I
RA region.",
RA Genomics 47:372-382(1998).
RA EMBL; Y14768; CAA75066.1; -;
RA GO: GO:0003793; P: defense/immunity protein activity; NAS.
RA GO: GO:0006955; P: immune response; NAS.
RA GO: GO:0006954; P: inflammatory response; NAS.
RA InterPro; IPR003599; IG-like.
RA InterPro; IPR007110; IG-like.
RA Pfam; PF00047; IG; 1.
RA SMART; SM00409; IG; 1.
RA PROSITE; PS50835; IG LIKE; 1.
RA SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;
SQ
Query Match 75.4%; Score 476.5; DB 4; Length 176;
Best Local Similarity 79.2%; Pred. No. 1.5e-43;
Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;
QY 1 LWSQPPPEIRTELGSSAFLPCSFNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
Db 19 LWSQPPPEIRTELGSSAFLPCSFNASQRLAIGSVTWFRDEVVPGKE----- 65
QY 61 PLASRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLVGVTGNGTRLVWEKEHPQLGAGT 120
Db 66 -----AELHIRDVRGHDASIYVCRVEVLGLVGVTGNGTRLVWEKEHPQLGAGT 113
RESULT 9
QBMJ00 PRELIMINARY; PRT; 115 AA.
ID QBMJ00
AC QBMJ00;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NKp30S.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RA "Molecular cloning of rhesus monkey NKp46 and NKp30 and identification
RA of NKp46SD and NKp30S.",
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RA EMBL; AY035217; AAKG3119.1; -;
RA InterPro; IPR003599; IG.
RA InterPro; IPR007110; IG-like.
RA Pfam; PF00047; IG; 1.
RA SMART; SM00409; IG; 1.
RA PROSITE; PS50835; IG LIKE; 1.
RA SEQUENCE 115 AA; 12871 MW; 1D34CC0B986DEB9F CRC64;
SQ
Query Match 73.7%; Score 466; DB 6; Length 115;
Best Local Similarity 94.6%; Pred. No. 1.2e-42;
Matches 88; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 LWSQPPPEIRTELGSSAFLPCSFNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
Db 19 LWSQPPPEIRTELGSSAFLPCSFNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASRFLHDHQAELHIRDVRGHDASIYVCRVE 93
Db 79 PLSSRFLRDHQAELHWDVRGHDAGIYVCRVE 111
RESULT 10
QBMJ01

```

ID QBMJ01 PRELIMINARY; PRT; 151 AA.
AC QBMJ01; 2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NKP30VLED.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
SEQUENCE FROM N.A.
RP LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification
of NKP46SD and NKP30S."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035216; AAK63118.1; -.
DR InterPro; IPR003599; IG.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;

Query Match 71.9%; Score 454.5; DB 6; Length 151;
Best Local Similarity 75.8%; Pred. No. 3.1e-41;
Matches 91; Conservative 1; Mismatches 3; Indels 25; Gaps 1;

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 66 -----AELHWDVRGHDAGIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 113

RESULT 11
Q8CFD9 PRELIMINARY; PRT; 192 AA.
AC Q8CFD9;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE lc7 protein precursor.
GN lc7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor lc7."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430418; CAD23066.1; -.
DR InterPro; IPR003599; IG.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;

Query Match 67.7%; Score 428; DB 11; Length 192;
Best Local Similarity 65.8%; Pred. No. 3e-38;
Matches 79; Conservative 19; Mismatches 22; Indels 0; Gaps 0;

```

```

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 79 SFSASQFIRGHKAGLLIQDISHDARIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138

RESULT 12
Q8QWM8 PRELIMINARY; PRT; 192 AA.
AC Q8QWM8;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NKP30.
GN NKP30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=LEW.
RC MEDLINE=22168131; PubMed=12180816;
RX Hsieh C.L., Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RA "NK cells and transplantation."
RL Transpl. Immunol. 9:111-114(2002).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=LEW.
RA Hsieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RT "Identification, Cloning, and Characterization of a Novel Rat NK
Receptor, rNKp30: a Molecule Expressed in Liver Allografts."
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY273824; AAP13457.1; -.
DR InterPro; IPR003599; IG.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

Query Match 67.1%; Score 424; DB 11; Length 192;
Best Local Similarity 67.8%; Pred. No. 8.2e-38;
Matches 78; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 115
DB 79 SFSASQFIRGHKAGLLIQDISHDARIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 133

RESULT 13
Q8CG11 PRELIMINARY; PRT; 192 AA.
AC Q8CG11;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NK receptor lc7 precursor.
GN lc7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=BN;

```

Search completed: February 26, 2004, 12:13:42
Job time : 52.2467 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 74.4828 Seconds
(without alignments)
455.215 Million cell updates/sec

Title: US-10-036-444-4

Perfect score: 632

Sequence: 1 LWSQPPETRLTSSSAFLP.....TNGTGLVVEKHPQLGAGT 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	120	4 AAE02771	Aae02771 Human Nkp
2	632	100.0	177	2 AAY06402	Aay06402 Human B-c
3	632	100.0	190	2 AAY06401	Aay06401 Human B-c
4	632	100.0	190	4 AAE02769	Aae02769 Human Nkp
5	632	100.0	201	2 AAY06403	Aay06403 Human B-c
6	617	97.6	135	5 AAE19109	Aae19109 Human Nkp
7	617	97.6	369	5 AAE19110	Aae19110 Human Nkp
8	118.5	18.8	139	5 AAU75565	Aau75565 Murine T
9	105.5	16.7	138	4 AAM24182	Aam24182 Rhesus mo
10	105	16.6	140	4 AAB68882	Aab68882 Human REC
11	104	16.5	262	2 AAR97726	Aar97726 B10 singl
12	103	16.3	136	4 ABB11287	Abb11287 Human mem
13	103	16.3	339	7 ADB65555	Adb65555 Human pro
14	103	16.3	3931	6 ABU07377	Abu07377 Human pro
15	101.5	16.1	267	2 AAW04300	Aaw04300 Murine T-
16	99	15.7	246	5 AAE20273	Aae20273 Human lun
17	99	15.7	270	1 AAP50256	Aap50256 Sequence
18	98.5	15.6	252	5 ABP45484	Abp45484 Human Bly
19	98.5	15.6	740	4 ABG28146	Abg28146 Novel hum
20	97.5	15.4	216	5 ABP58185	Abp58185 CRLA-4-GC
21	96.5	15.3	125	5 ABP2181	Abp2181 Human mon
22	96.5	15.3	350	4 ABB69289	Abb69289 Drosophil
23	95.5	15.1	131	6 AAO39772	Aao39772 Rat myeli
24	95.5	15.1	150	2 AAW97817	Aaw97817 Rat myeli
25	94	14.9	526	2 AAW97814	Aaw97814 Human but

26	94	14.9	526	5 AAO15804	Aao15804 Human but
27	92.5	14.6	235	4 AAG64474	Aag64474 Human typ
28	92.5	14.6	388	5 ABB07681	Abb07681 MOG-Pc-fu
29	92.5	14.6	388	6 ADA14289	Ada14289 Mutated M
30	92.5	14.6	388	6 ADA14265	Ada14265 Human imm
31	92	14.6	174	3 AAB08208	Aab08208 Amino aci
32	91.5	14.5	223	2 AAY41083	Aay41083 Canine CT
33	91.5	14.5	268	2 AAR77288	Aar77288 T-cell re
34	91.5	14.5	269	3 AAY69995	Aay69995 Human rec
35	91	14.4	205	2 AAY41173	Aay41173 Llana Vhh
36	91	14.4	253	5 ABP45534	Abp45534 Human Bly
37	90.5	14.3	235	4 AAG64476	Aag64476 Human typ
38	90.5	14.3	502	6 ABJ37109	Abj37109 Concatame
39	89.5	14.2	110	6 ABR55818	Abr55818 Lambda ch
40	89.5	14.2	137	3 AAB08207	Aab08207 cDNA enco
41	89.5	14.2	383	5 ABB78103	Abb78103 Amino aci
42	89.5	14.2	383	5 AAU75121	Aau75121 Human sol
43	89.5	14.2	383	6 ABP56719	Abp56719 CTLM4 mut
44	89.5	14.2	383	7 ADD89007	Add89007 L104EA29L
45	89.5	14.2	391	2 AAR97659	Aar97659 Single ch

ALIGNMENTS

RESULT 1

AAE02771
ID AAE02771 standard; protein; 120 AA.

XX AAE02771;

XX AC

XX DT 06-AUG-2001 (first entry)

XX DE Human Nkp30 receptor extracellular region sequence.

XX KW Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;

XX KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

XX KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

XX KW therapy; extracellular region.

XX OS Homo sapiens.

XX PN WO200136630-A2.

XX PD 25-MAY-2001.

XX PF 15-NOV-2000; 2000MO-EP011697.

XX PR 15-NOV-1999; 99CA-02288307.

XX PA (INNA-) INNATE PHARMA SAS.

XX PA (UYGE-) UNIV GENOVA.

XX PI Moretta A, Bottino C, Biassoni R;

XX WPI; 2001-329221/34.

XX Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.

XX Claim 1; Fig 7B; 83pp; English.

XX The invention relates to human Nkp30 receptor and its corresponding cDNA

XX molecule which is involved in natural cytotoxicity mediated by natural

XX killer (NK) cells and antibodies that identify the same Nkp30 receptor

XX is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively

XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are

XX useful for detecting and/or quantifying the presence of NK cells in a

XX biological sample. The invention also provide kits for detecting and/or

XX quantifying the presence of NK cells, for the selective removal of NK

XX cells from a biological sample, for the positive and selective

XX purification of NK cells from a biological sample and for the in vitro

stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GVT) and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. Nkp30 antibodies are useful for identifying Nkp30 natural ligands and allow assessment of the level of surface Nkp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence Nkp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the extracellular region of human Nkp30 receptor

XX Sequence 120 AA;

Query Match 100.0%; Score 632; DB 4; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.6e-62;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
Db 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
Db 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120

RESULT 2

AY06402
ID AAY06402 standard; protein; 177 AA.

XX AAY06402;

DT 20-SEP-1999 (first entry)

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.

XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
KW signal transduction; immunomodulator; antiinflammatory;
KW autoimmune disease; inflammation; gene therapy; diagnosis.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..12

FT /note= "leader peptide"

FT Protein 13..177

FT /note= "mature protein"

FT Modified-site 42

FT /note= "N-glycosylated"

FT Modified-site 68

FT /note= "N-glycosylated"

FT Modified-site 121

FT /note= "N-glycosylated"

FT Domain 139..162

FT /note= "transmembrane domain"

FT Peptide 166..177

FT /note= "alternatively spliced C-terminal end"

XX WO9923867-A2.

XX 20-MAY-1999.

XX 05-NOV-1998; 98WO-US023826.

XX 07-NOV-1997; 97US-0064761P.

XX (BIOJ) BIOGEN INC.

XX Browning J;

XX WPI; 1999-418423/35.

DR N-PSDB; AAX59348.
XX Novel B-cell myelin oligodendrocyte glycoproteins.
PT Claim 2; Page 43; 43pp; English.

CC This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

XX Sequence 177 AA;

Query Match 100.0%; Score 632; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
Db 19 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78

QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
Db 79 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138

RESULT 3

AY06401
ID AAY06401 standard; protein; 190 AA.

XX AAY06401;

DT 20-SEP-1999 (first entry)

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.

XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
KW signal transduction; immunomodulator; antiinflammatory;
KW autoimmune disease; inflammation; gene therapy; diagnosis.

OS Homo sapiens.

XX Key Location/Qualifiers
FH Peptide 1..12

FT /note= "leader peptide"

FT Protein 13..190

FT /note= "mature protein"

FT Modified-site 42

FT /note= "N-glycosylated"

FT Modified-site 68

FT /note= "N-glycosylated"

FT Modified-site 121

FT /note= "N-glycosylated"

FT Domain 139..162

FT /note= "transmembrane domain"

FT Peptide 166..190

FT /note= "alternatively spliced C-terminal end"

XX WO9923867-A2.

XX 20-MAY-1999.

XX 05-NOV-1998; 98WO-US023826.

XX 07-NOV-1997; 97US-0064761P.
XX (BIOJ) BIOGEN INC.
XX Browning J;
XX WPI; 1999-418423/35.
XX N-PSDB; AAX59347.
XX Novel B-cell myelin oligodendrocyte glycoproteins.
XX Claim 2; Page 42; 43pp; English.
XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)
XX SQ Sequence 190 AA;
Query Match 100.0%; Score 632; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.8e-62; Indels 0; Gaps 0;
Matches 120; Conservative 0; Mismatches 0;
Qy 1 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 60
Db 19 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 78
Qy 61 PLASSRFLHDHQAELHRLDVRGHDASIVYCRVEVLGVGTGTGTRLVVKEHPQLGAGT 120
Db 79 PLASSRFLHDHQAELHRLDVRGHDASIVYCRVEVLGVGTGTGTRLVVKEHPQLGAGT 138
RESULT 4
AAE02769
ID AAE02769 standard; protein; 190 AA.
AC AAE02769;
XX 06-AUG-2001 (first entry)
XX Human Nkp30 receptor.
XX Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;
XX melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
XX immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
XX therapy.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Peptide 1..18
XX Protein /label= Signal_peptide
XX Region /label= Mature_Nkp30_receptor_protein
XX /label= Extracellular_region
XX /note= "Forms an immunoglobulin (Ig) V-like domain"
XX Modified-site 42
XX Modified-site /note= "N-glycosylation site"
XX 121
XX /note= "N-glycosylation site"

FT Region 139..157
FT /label= Transmembrane_region
FT 158..190
FT /label= Intracellular_region
XX WO200136630-A2.
XX 25-MAY-2001.
XX 15-NOV-2000; 2000WO-EP011697.
XX 15-NOV-1999; 99CA-02288307.
XX 15-NOV-1999; 99US-00440514.
XX (INNA-) INNATE PHARMA SAS.
XX (UYGE-) UNIV GENOVA.
XX Moretta A, Bottino C, Biassoni R;
XX WPI; 2001-329221/34.
XX N-PSDB; AAD06564.
XX Novel compound, useful for detection and/or quantifying the presence of
XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.
XX Claim 1; Fig 7B; 83pp; English.
XX The invention relates to human Nkp30 receptor and its corresponding cDNA
XX molecule which is involved in natural cytotoxicity mediated by natural
XX killer (NK) cells and antibodies that identify the same. Nkp30 receptor
XX is a member of immunoglobulin super family (Ig-Sp). Nkp30 is selectively
XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are
XX useful for detecting and/or quantifying the presence of NK cells in a
XX biological sample. The invention also provide kits for detecting and/or
XX quantifying the presence of NK cells, for the selective removal of NK
XX cells from a biological sample, for the positive and selective
XX purification of NK cells from a biological sample and for the in vitro
XX stimulation of NK cell cytotoxicity. The invention further provides a
XX pharmaceutical composition which is used as a drug for grafting
XX enhancement, graft versus host (GVH) inhibition, stimulation of graft
XX versus tumour (Gvt) and especially graft versus leukaemia (GvL), and for
XX the prevention, palliation and/or therapy of solid or liquid tumours,
XX such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
XX microorganism, notably viral infection. Nkp30 antibodies are useful for
XX identifying Nkp30 natural ligands and allow assessment of the level of
XX surface Nkp30 ligand expressed on an NK-susceptible target cell and the
XX comparison of this level to the standard physiological one. Hence Nkp30
XX antibodies are useful in the diagnosis of tumours or of infection. The
XX present sequence is human Nkp30 receptor
XX SQ Sequence 190 AA;
Query Match 100.0%; Score 632; DB 4; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.8e-62; Indels 0; Gaps 0;
Matches 120; Conservative 0; Mismatches 0;
Qy 1 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 60
Db 19 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 78
Qy 61 PLASSRFLHDHQAELHRLDVRGHDASIVYCRVEVLGVGTGTGTRLVVKEHPQLGAGT 120
Db 79 PLASSRFLHDHQAELHRLDVRGHDASIVYCRVEVLGVGTGTGTRLVVKEHPQLGAGT 138
RESULT 5
AAV06403
ID AAV06403 standard; protein; 201 AA.
XX AAV06403;
XX 20-SEP-1999 (first entry)
XX

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 XX MBOG; B-cell myelin oligodendrocyte glycoprotein; human;
 KW signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 XX Peptide 1..12
 FT /note= "leader peptide"
 FT Protein 13..201
 FT /note= "mature protein"
 FT Modified-site 42
 FT /note= "N-glycosylated"
 FT Modified-site 68
 FT /note= "N-glycosylated"
 FT Modified-site 121
 FT /note= "N-glycosylated"
 FT Domain 139..162
 FT /note= "transmembrane domain"
 FT Peptide 166..201
 FT /note= "alternatively spliced C-terminal end"
 XX
 PN WO9923867-A2.
 XX
 PD 20-MAY-1999.
 XX
 PF 05-NOV-1998; 98WO-US023826.
 XX
 PR 07-NOV-1997; 97US-0064761P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Browning J;
 XX
 DR WPI; 1999-418423/35.
 DR N-PSDB; AAX59349.
 XX
 PT Novel B-cell myelin oligodendrocyte glycoproteins.
 XX
 PS Claim 2; Page 43; 43pp; English.
 XX
 CC This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AAX06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 SQ Sequence 201 AA;
 Query Match 100.0%; Score 632; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 3e-62;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDEVPGKEVRNGTPEFRGLA 60
 DB 19 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDEVPGKEVRNGTPEFRGLA 78
 QY 61 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 120
 DB 79 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 138

RESULT 6
 AAE19109
 ID AAE19109 standard; protein; 135 AA.
 XX
 AC AAE19109;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human Nkp30 protein.
 XX
 KW Human; natural killer cell activating protein; Nkp46; therapy; virucide;
 KW viral infection; natural killer cell; NK; Nkp44; imaging agent; cancer;
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; Nkp30.
 XX
 OS Homo sapiens.
 XX
 PN WO200208287-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-IL000664.
 XX
 PR 20-JUL-2000; 2000IL-00137419.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA (UTNE) UNIV BEN-GURION NEGEV.
 XX
 PI Mandelboim O, Porgador A;
 XX
 DR WPI; 2002-195870/25.
 DR N-PSDB; AAD30466.
 XX
 PT New targeting complex capable of targeting an active substance to a
 PT target cell, comprising a target recognition segment and an active
 PT segment, useful for treating pathologies associated with viral infections
 PT or cancer.
 XX
 PS Example 1; Page 108; 113pp; English.
 XX
 CC The invention relates to compositions and methods for the treatment and
 CC detection of a variety of viral infections, by using complex agents
 CC comprising the natural killer (NK) cells activating proteins, Nkp46 and
 CC Nkp44 and functional fragments thereof, linked to therapeutic or imaging
 CC agents. The complex is useful for treating pathologies associated with
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECMV, MVM or herpes virus)
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
 CC the imaging and monitoring of cancer. The complex may also be used to
 CC detect the presence of abnormal cells in a sample. The antibodies can be
 CC used to qualitatively or quantitatively detect the ligand for the
 CC complex. The present sequence is human Nkp30 protein
 XX
 SQ Sequence 135 AA;
 Query Match 97.6%; Score 617; DB 5; Length 135;
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDEVPGKEVRNGTPEFRGLA 60
 DB 19 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDEVPGKEVRNGTPEFRGLA 78
 QY 61 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 117
 DB 79 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 135

RESULT 7
 AAE19110
 ID AAE19110 standard; protein; 369 AA.
 XX
 AC AAE19110;
 XX

29-AUG-2003 (revised)
21-MAY-2002 (first entry)
Human NKp30-IgG fusion protein.
Human, natural killer cell activating protein; NKp46; therapy; virucide;
viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;
detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;
immunoglobulin G; fusion protein.
Homo sapiens.
Chimeric.
Key Location/Qualifiers
Region 1..135
/note= "Human NKp30"
Region 136..369
/note= "Human IgG"
WO200208287-A2.
31-JAN-2002.
19-JUL-2001; 2001WO-IL000664.
20-JUL-2000; 2000IL-00137419.
(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
(UTNE) UNIV BEN-GURION NEGEV.
Mandelbeim O, Porzador A;
WPI; 2002-195870/25.
N-PSDB; AAD30467.
New targeting complex capable of targeting an active substance to a
target cell, comprising a target recognition segment and an active
segment, useful for treating pathologies associated with viral infections
or cancer.
Example 1; Page 108-110; 113pp; English.
The invention relates to compositions and methods for the treatment and
detection of a variety of viral infections, by using complex agents
comprising the natural killer (NK) cells activating proteins, NKp46 and
NKp44 and functional fragments thereof, linked to therapeutic or imaging
agents. The complex is useful for treating pathologies associated with
viral infections (e.g. infections caused by influenza virus, HIV, Epstein
-Barr virus, cytomegalovirus, vaccinia virus, EMCV, MVM or herpes virus)
and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
the imaging and monitoring of cancer. The complex may also be used to
detect the presence of abnormal cells in a sample. The antibodies can be
used to qualitatively or quantitatively detect the ligand for the
complex. The present sequence is human NKp30- immunoglobulin G (IgG) Fc
region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)
Sequence 369 AA;

Query Match 97.6%; Score 617; DB 5; Length 369;
Best Local Similarity 100.0%; Pred. No. 3e-60;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LWSVQPEIRTELGSSAFPLPCSNASQGRLAGSVTWFRDEVPVPGKEVNGTPFRGRILA 60
Db 19 LWSVQPEIRTELGSSAFPLPCSNASQGRLAGSVTWFRDEVPVPGKEVNGTPFRGRILA 78
QY 61 PLASSRFLHDHQAELHIRDVRGHDSAIYVCRVEVLGLVGTGNGRLVVEKEHPQLG 117
Db 79 PLASSRFLHDHQAELHIRDVRGHDSAIYVCRVEVLGLVGTGNGRLVVEKEHPQLG 135

RESULT 8
AAU75565

ID AC AAU75565 standard; protein; 139 AA.
AC AAU75565;
DT 23-APR-2002 (first entry)
DE Murine T cell receptor beta chain.
KW Immunoglobulin superfamily; GP286; human; immunosuppressive;
autoimmune disease; rheumatoid arthritis; cancer; multiple sclerosis;
acquired immune deficiency syndrome; AIDS; inflammatory disorder;
pancreatitis; antirheumatic; antipsoriatic; dermatologic; antianaemic;
cytostatic; antileukaemic; antiasthmatic; antiallergic;
T cell receptor beta chain.
OS Mus sp.
XX WO200200727-A2.
PN 03-JAN-2002.
PD 22-JUN-2001; 2001WO-US020038.
PF 23-JUN-2000; 2000US-0213630P.
PR 13-APR-2001; 2001US-0283813P.
XX (BIOJ) BIOGEN INC.
PA Carulli JP, Lukashin AV, Kilburn DR, Mathur P;
PI WPI; 2002-090520/12.
XX Isolated polynucleotide encoding a novel human immunoglobulin superfamily
member, named GP286, useful in the treatment of a disease condition that
relates to the immune system, e.g. a transplantation disorder or an
autoimmune disease.
PS Disclosure; Page 167; 180pp; English.
XX The invention relates to an isolated polynucleotide (I) encoding a novel
human immunoglobulin superfamily member, named GP286. The GP286
polynucleotide and polypeptide are useful for the treatment of a disease
condition that relates to the immune system, preferably a disease
condition that relates to T cells, e.g. a transplantation disorder, an
autoimmune disease (e.g. rheumatoid arthritis, systemic lupus
erythematosus, psoriasis, Sjogren's Syndrome, thyroiditis, Graves'
disease, pulmonary fibrosis, bronchiolitis obliterans, haemolytic anaemia
or Wegener's granulomatosis), cancer (e.g. leukaemia or lymphoma),
multiple sclerosis, graft versus host disease, Kawasaki syndrome,
acquired immune deficiency syndrome (e.g. AIDS) or an inflammatory
disorder (e.g. asthma, allergies, adult respiratory distress syndrome and
acute pancreatitis or chronic pancreatitis). The GP286 nucleic acids can
be used as probes for detecting, characterising and quantifying GP286
nucleic acids. The present sequence represents the amino acid sequence of
murine T cell receptor beta chain used in the method of the invention
Sequence 139 AA;
Query Match 18.8%; Score 118.5; DB 5; Length 139;
Best Local Similarity 30.2%; Pred. No. 4.6e-05;
Matches 39; Conservative 17; Mismatches 40; Indels 33; Gaps 7;
QY 2 WVS-----OPPRTLEGSSAFPLPCSNASQGRLAGSVTWFRDE-----VWPG 45
Db 18 WVSQDQVKQSPSALSLOEGTSSALRCNFS-----IATTTVQVFLQNSRGLMFLVLP- 71
QY 46 KEVNGTPFRGRILAPLASSRFLHDHQAELHIRDVRGHDSAIYVCRVEVLGLG--VGTGN 103
Db 72 -----GTYE-NGRLLKSTFNK---ESYSLTHIRDAQLSDSGTFCRAEVEGTGSKLSFGK 122
QY 104 GTRLVVEKE 112
Db 123 GAKLTVSPD 131

CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention
XX

XX Sequence 136 AA;

Query Match 16.3%; Score 103; DB 4; Length 136;
Best Local Similarity 28.7%; Pred. No. 0.0024;
Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;

QY 7 PEIRTEGSSAFPCSFNASQRLAIGSVTFWRDEVVGKVRNGTFEFGRLAPLASSR 66

Db 32 PDLSTTEGSHAFPCCKARGSP-----PNTWKD-----GQVSGAEGK 71

QY 67 FLHDHQAEHLHIRDYRGHDASIVYCVRE 93

Db 72 FTIQPSGELLVKNLEGGDAGTYTCTAE 98

RESULT 13

ADBE6555

ID ADB65555 standard; protein; 329 AA.

AC ADB65555;

DT 04-DEC-2003 (first entry)

XX Human protein encoded by clone THYM20028150.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.

XX Homo sapiens.

XX EP1308459-A2.

PD 07-MAY-2003.

PF 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

PR 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-450961/43.

DR N-PSDB; ADB63585.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
PI marker or medicines for regulation of their expression and activity, or

PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesising the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the invention. Note: Some of the
CC sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.

XX Sequence 329 AA;

Query Match 16.3%; Score 103; DB 7; Length 329;

Best Local Similarity 28.7%; Pred. No. 0.0074;

Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;

QY 7 PEIRTEGSSAFPCSFNASQRLAIGSVTFWRDEVVGKVRNGTFEFGRLAPLASSR 66

Db 176 PDLSTTEGSHAFPCCKARGSP-----PNTWKD-----GQVSGAEGK 215

QY 67 FLHDHQAEHLHIRDYRGHDASIVYCVRE 93

Db 216 FTIQPSGELLVKNLEGGDAGTYTCTAE 242

RESULT 14

ABU07377

ID ABU07377 standard; protein; 3931 AA.

XX AC ABU07377;

XX 28-JAN-2003 (first entry)

XX Human protein NOV9.

XX Human; NOV9; cardiomyopathy; atherosclerosis; cancer; hypertension;
KW diabetes; inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin A nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW Parkinson's disease; Goitre; infection; stroke; muscular dystrophy;
KW epilepsy; wasting disorder; neurogenesis; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
KW Gene therapy; single nucleotide polymorphism; SNP.

XX Homo sapiens.

XX WO200285922-A2.

XX 31-OCT-2002.

XX 11-APR-2002; 2002WO-US011634.

XX 23-APR-2001; 2001US-0285748P.

PR 24-APR-2001; 2001US-0286068P.


```

PR 25-APR-2001; 2001US-0286292P.
PR 03-MAY-2001; 2001US-0288334P.
PR 16-MAY-2001; 2001US-0291241P.
PR 14-SEP-2001; 2001US-0322284P.
XX (CURA-) CURAGEN CORP.
FA (MILL-) MILLENNIUM PHARM INC.
XX
PI Pena CEA, Guo X, Shinkets RA, Padigaru M, Kekuda R, Spytek KA;
PI Mehraban F, Topper JN, Malyankar UM, Wasserman S, Edinger S;
PI Smithson G, Gunther E, Komuves L;
XX
DR WPI; 2003-058712/05.
DR N-PSDB; ABX10231.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
FS Claim 1; Page 65-66; 301pp; English.
XX
XX The invention relates to an isolated polypeptide comprising any one of 17
CC human NOVX (1-9, 10a, 10b, 11-16) appearing as ABU07369-ABU07385, a
CC mature form of it, or a variant of them, where one or more residues of
CC the variant differs in not more than 15 % from the residues of the
CC sequence of them and their encoding polynucleotides appearing as ABX10223
CC -ABX10239. Also included are NOVX expression vectors, transformed cells,
CC antibodies, identifying an agent that binds to or modulates the
CC expression or activity of NOVX and screening for a modulator of activity
CC or of latency or predilection to a NOVX-associated disorder. The NOVX
CC polypeptides, polynucleotides and antibodies are useful in manufacturing
CC a medicament for treating or preventing a syndrome associated with NOVX-
CC associated disorder, such as cardiomyopathy, atherosclerosis, cancer,
CC hyperension, diabetes, inflammation, autoimmune disorders, allergies,
CC blood disorders, acquired immunodeficiency syndrome (AIDS), obesity,
CC asthma, immunoglobulin (IgA) nephropathy, cirrhosis, arthritis,
CC Alzheimer's disease, Parkinson's disease, goitre, infections (e.g.
CC bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy, and
CC other wasting disorders associated with chronic diseases. The nucleic
CC acids and polypeptides may also be used as targets for the identification
CC of small molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids are further used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine, and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence represents a NOVX protein
XX
SQ Sequence 3931 AA;
Query Match 16.3%; Score 103; DB 6; Length 3931;
Best Local Similarity 28.7%; Pred No. 0.17;
Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;
QY 7 PEIRTEGSAFLPCSFNASQGRLAIGSVTFWDEVPVGVKEVRNGTPEFRGLAPLASSR 66
Db 2759 PDLSTTEGSHAFPCKARGSP-----PNTWDXD-----GQPVSGAEGK 2798
QY 67 FLHDHQAEHLHVRDVRGHDASIVYCRVE 93
Db 2799 FTIQPSGELLVKNLEGGDAGTYTCTAE 2825
RESULT 15
AAW04300
XX AAW04300 standard; protein; 267 AA.
AC AAW04300;
XX
XX 02-JAN-1997 (first entry)
XX

```

```

DE Murine T-cell receptor alpha chain.
XX
XX Antigen-specific glycosylation inhibiting factor; AGGIF; lymphokine;
KW T-cell receptor; TCR; immunosuppressant; immunosuppressive;
KW hypersensitivity; graft rejection; autoimmune disease.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FT Peptide 1..20
FT Region /label= Leader_peptide
FT Region 21..109
FT Region /label= V-alpha_region
FT Region 110..130
FT Region /label= J-alpha_region
FT Region 131..267
FT Region /label= C-alpha_region
XX
XX WO9631617-A1.
XX
XX 10-OCT-1996.
XX
XX 25-SEP-1995; 95WO-US012309.
XX
XX 04-APR-1995; 95US-00416336.
XX
XX (LJOL-) LA JOLLA INST ALLERGY & IMMUNOLOGY.
XX
XX Ishizaka K, Ishii Y;
XX
XX WPI; 1996-465037/46.
XX N-PSDB; AAT38389.
XX
XX Recombinant antigen-specific glycosylation-inhibiting factor - useful as
XX immunosuppressant.
XX
XX Example 3; Fig 2A; 72pp; English.
XX
XX A cDNA clone (AAT38389) coding for the T-cell receptor (TCR) alpha chain
XX (AAW04300) was isolated from T-cell hybridoma 23F1 cells (ovaalbumin-
XX specific). TCR alpha plays a role in the formation of antigen-specific
XX glycosylation inhibiting factor (AGGIF) chains which directly bind to an
XX antigen and suppress the immune response generated against that antigen;
XX AGGIF is an expression product of the TCR alpha chain gene (see also
XX AAT38389). A polynucleotide encoding AGGIF can be obt'd. from activated
XX suppressor T-cells using a probe based on non-specific GIP cDNA (see also
XX AAT38390-91) and a probe based on TCR alpha cDNA, and used to direct
XX expression of Aggif in host cells
XX
XX Sequence 267 AA;
Query Match 16.1%; Score 101.5; DB 2; Length 267;
Best Local Similarity 30.2%; Pred No. 0.0083;
Matches 38; Conservative 16; Mismatches 39; Indels 33; Gaps 8;
QY 2 WVS-----QPPEIRTEGSAFLPCSFNASQGRLAIGSVTFWDEVPVGVKEVRN----- 50
Db 18 WVGSGKVKQSPSALSQEGTNSALRCNFS-----IAATTYQWFQFN--PRGSLMNLFLV 70
QY 51 -GTPEFRGLAPLASSRFLHDHQAEHLHVRDVRGHDASIVYCRVEVLGLVGT-----GN 103
Db 71 PGTKE-NGRLKSTFNSK---ESYSTLHIRDAQLEDSTGYTCAAE-----GGGSNYKLTFGK 122
QY 104 GTRLVV 109
Db 123 GTLLTV 128

```

Search completed: February 26, 2004, 12:10:06
Job time : 77.6828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 41.6976 Seconds
(without alignments)
607.670 Million cell updates/sec

Title: US-10-036-444-4
Perfect score: 632
Sequence: 1 LWVSQPEIRTEGSSAFPLP.....TGNTRLVWEKHPOLGAGT 120

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	120	13	US-10-036-444-4
2	632	100.0	190	13	Sequence 4, Appli
3	118.5	18.8	139	14	Sequence 2, Appli
4	103	16.3	329	15	Sequence 6, Appli
5	103	16.3	329	15	Sequence 3709, Ap
6	99	15.7	246	10	Sequence 18, Appl
7	98.5	15.6	252	10	Sequence 49, Appl
8	94	14.9	526	9	Sequence 1495, Ap
9	94	14.9	526	9	Sequence 13, Appli
10	94	14.9	526	9	Sequence 19, Appl
11	94	14.9	526	14	Sequence 7, Appli
12	93	14.7	1598	10	Sequence 59, Appl
13	93	14.7	1694	10	Sequence 57, Appl
14	91	14.4	205	9	Sequence 9, Appli
15	91	14.4	253	10	Sequence 1545, Ap

ALIGNMENTS

RESULT 1

US-10-036-444-4
; Sequence 4, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-PR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-4

Query Match 100.0%; Score 632; DB 13; Length 120;
Best Local Similarity 100.0%; Pred. No. 8e-62;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LWVSQPEIRTEGSSAFPLP...CSFNASQGR...LAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
Db 1 LWVSQPEIRTEGSSAFPLP...CSFNASQGR...LAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
QY 61 PLASRFLHDHQAELHLDVRGHDASIVYCRVEVLGLGVGTGCTRLVWEKEHOLGAGT 120
Db 61 PLASRFLHDHQAELHLDVRGHDASIVYCRVEVLGLGVGTGCTRLVWEKEHOLGAGT 120

```
RESULT 2
US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; TITLE OF INVENTION: "NO. US20020142445A1 triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-PR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; PRIOR FILING DATE: 2002-01-07
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-2

Query Match      100.0%; Score 632; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.4e-61;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWVSQPPPIRTLEGSSAFPLPCSFNASQGRLAIGSVTWFDRVWPGKEVRNGTPEFRGLA 60
DB 19 LWVSQPPPIRTLEGSSAFPLPCSFNASQGRLAIGSVTWFDRVWPGKEVRNGTPEFRGLA 78

QY 61 PLASRFLLHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGCTRLVWKEHPOLGAST 120
DB 79 PLASRFLLHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGCTRLVWKEHPOLGAGT 138

RESULT 3
US-10-312-495-6
; Sequence 6, Application US/10312495
; Publication No. US20030165495A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A099PCT000454-124
; CURRENT APPLICATION NUMBER: US/10/312.495
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/213,630
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/283,813
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-312-495-6

Query Match      18.8%; Score 118.5; DB 14; Length 139;
Best Local Similarity 30.2%; Pred. No. 3.4e-05;
Matches 39; Conservative 17; Mismatches 40; Indels 33; Gaps 7;

QY 2 WVS-----OPPEITLGGSAFLPCSFNASQGRLAIGSVTWFDR-----VVP 45
DB 18 WVSQGVQKSPALSISQETSSALRCNFS-----IATTVQWFLQNSRGLMNLFLYLP- 71
QY 46 KEVRNGTPEFRGLAPLASSRFLHQAELHIRDVRGHDAIYVCRVEVLGLG--VGTGN 103
DB 72 -----GKE--NRLKSTFNSK---ESYTLHIRDAQLESDSGTYFCAAEVGTGSKLSFGK 122
QY 104 GTRLVVEKE 112
```

```
DB 123 GAKLTVSPD 131

RESULT 4
US-10-104-047-3709
; Sequence 3709, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104.047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3709
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3709

Query Match      16.3%; Score 103; DB 15; Length 329;
Best Local Similarity 28.7%; Pred. No. 0.005;
Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;

QY 7 PEIRTEGSAFLPCSFNASQGRLAIGSVTWFDRVWPGKEVRNGTPEFRGLAPLASSR 66
DB 176 PDLSTTEGSHAPLPCKARGSP-----PNTWDK-----QOPVSGAGK 215

QY 67 FLHDHQAELHIRDVRGHDAIYVCRVE 93
DB 216 FTIQSGELLVKNLEGQDAGTYTCTAE 242

RESULT 5
US-10-120-801-18
; Sequence 18, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120.801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
```

```

; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 3931
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-120-801-18

Query Match      16.3%; Score 103; DB 15; Length 3931;
Best Local Similarity 28.7%; Pred. No. 0.11; 31; Indels 20; Gaps 2;
Matches 25; Conservative 11; Mismatches 11;

QY 7 PEIRTEGSSAFPCSNASQGRLAIGSVTFRDVEVPGKEVRNGTPEFRGLAPLASSR 66
DQ 2759 PDLSTTEGSHAFPCSKARGSP-----PNIITWDK-----GQVSGAECK 2798
DQ 67 FLHDHQELHVRDVRGHDAISYVCRVE 93
DQ 2799 FTIQPSGELLVKNLEGGQAGTYTCTAE 2825

RESULT 6
US-09-909-567B-49
; Sequence 49, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DEX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-49

Query Match      15.7%; Score 99; DB 10; Length 246;
Best Local Similarity 30.5%; Pred. No. 0.0096;
Matches 39; Conservative 11; Mismatches 42; Indels 36; Gaps 7;

QY 3 VSQPPERTLEGSSAFPCSNASQGRLAIGSVTFRDVEVPGKEVR-----NGTPEF 55
DQ 23 LTQPPSASALGASVTLTCSVSSDYKNL---EVDWFQQR---PGKGRFVRVWVGTVGVGF 77
QY 56 RGLAP-----LASSRFLHDHQELHVRDVRGHDAISYVCRVEVLGLVGT----- 101
DQ 78 RADIPDRPSVSGSLNRF-----LTIRNIEDESDYHCGTD---LGSGTFSVSWVF 127

QY 102 GNGTRLVV 109
DQ 128 GGGTKLTV 135

RESULT 7
US-09-880-748-1495
; Sequence 1495, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1495
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1495

Query Match      15.6%; Score 98.5; DB 10; Length 252;
Best Local Similarity 28.8%; Pred. No. 0.011;
Matches 34; Conservative 16; Mismatches 45; Indels 23; Gaps 6;

QY 3 VSQPPERTLEGSSAFPCSNASQGRLAIGSVTFRDVEVPGKEVR-----NGT 52
DQ 145 VTQPPSVSAAPQQRVTISCS--GSSSNIGINSVSYQQ--LFGKAPKLLIYNNNRPSGV 200
QY 53 PEFRGRGLAPLASSRFLHDHQELHVRDVRGHDAISYVCRVEVLGL--GVGTGNGTRLVV 109
DQ 201 P---GRFSGSKSGT-----SASLAITGLQAEDEADYICQSYDNSLNGVLFGGGTQLTV 250

RESULT 8
US-09-910-174A-9
; Sequence 9, Application US/09910174A
; Patent No. US20020106730A1
; GENERAL INFORMATION:
; APPLICANT: Coyle, Anthony J.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Manning, Stephen
; TITLE OF INVENTION: B7-H2 Molecules, No. US20020106730A1 Members of the B7
; TITLE OF INVENTION: Family and Uses Thereof
; FILE REFERENCE: 35800/236924
; CURRENT APPLICATION NUMBER: US/09/910,174A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/620,461
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-910-174A-9

Query Match      14.9%; Score 94; DB 9; Length 526;
Best Local Similarity 39.6%; Pred. No. 0.088;
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;

QY 6 PPE-INTLEGSSAFPC--SFNASQGRLAIGSVTFRDVEVPGKEVR-RNG-----TPE 54
DQ 34 PPEPILAVVGEDAELPCRLSPNASEHL---ELRFRKKVSPAVLVHRDGRGQEAQOMPE 90
QY 55 FRGLAPLASSRFLHDHQELHVRDVRGHDAISYVCR 90
DQ 91 YRGR-ATLVQDGIAGRVA-LRIGRVVSDDGEYTC 124

RESULT 9
US-09-955-866-13
; Sequence 13, Application US/09955866
; Patent No. US20020107363A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Michael
; APPLICANT: Sullivan, John K.
; APPLICANT: Holst, Paige

```


; PRIOR APPLICATION NUMBER: 60/263,688
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/263,694
 ; PRIOR FILING DATE: 2001-01-24
 ; NUMBER OF SEQ ID NOS: 155
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 59
 ; LENGTH: 1598
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-863-776-59

Query Match 14.7%; Score 93; DB 10; Length 1598;
 Best Local Similarity 29.4%; Pred. No. 0.45; 34; Indels 26; Gaps 8;
 Matches 32; Conservative 17; Mismatches 17;

QY 2 W-VSQPEIRTEGSSAFLPCSFN-----ASQRLAIGSVTWFRDEVPGKE---VRNG 51
 DB 21 WGVSSPKNVQGLSGSCLLIFCIPSPADVPVNSGITAI-----WYD--YSGKQVVIHSG 74
 QY 52 TPE-----FRGLAPLASSRFLHDHQ-AELHIRDVRGHDASIVYCRVEV 94
 DB 75 DPKLVDRKFRGRaelMGNM-----DHKVCNLLKDLKPEDSGTYNFRFEI 119

RESULT 13
 US-09-863-776-57
 ; Sequence 57, Application US/09863776
 ; Publication No. US20030198953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Spytex, Kimberly A
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Mishra, Vishnu
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Li, Li
 ; APPLICANT: Taupier, Raymond J
 ; APPLICANT: Gangolli, Esha
 ; TITLE OF INVENTION: No. US20030198953A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-020
 ; CURRENT APPLICATION NUMBER: US/09/863,776
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: 09/540,763
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: 60/206,679
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 60/206,688
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 60/206,829
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 60/207,748
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: 60/207,798
 ; PRIOR FILING DATE: 2000-05-30
 ; PRIOR APPLICATION NUMBER: 60/208,263
 ; PRIOR FILING DATE: 2000-05-31
 ; PRIOR APPLICATION NUMBER: 60/208,831
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: 60/209,451
 ; PRIOR FILING DATE: 2000-06-05
 ; PRIOR APPLICATION NUMBER: 60/210,060
 ; PRIOR FILING DATE: 2000-06-07
 ; PRIOR APPLICATION NUMBER: 60/219,507
 ; PRIOR FILING DATE: 2000-07-20
 ; PRIOR APPLICATION NUMBER: 60/221,337
 ; PRIOR FILING DATE: 2000-07-26
 ; PRIOR APPLICATION NUMBER: 60/221,927
 ; PRIOR FILING DATE: 2000-07-31
 ; PRIOR APPLICATION NUMBER: 60/263,135
 ; PRIOR FILING DATE: 2001-01-19

; PRIOR APPLICATION NUMBER: 60/263,688
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/263,694
 ; PRIOR FILING DATE: 2001-01-24
 ; NUMBER OF SEQ ID NOS: 155
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 57
 ; LENGTH: 1694
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-863-776-57

Query Match 14.7%; Score 93; DB 10; Length 1694;
 Best Local Similarity 29.4%; Pred. No. 0.48; 34; Indels 26; Gaps 8;
 Matches 32; Conservative 17; Mismatches 17;

QY 2 W-VSQPEIRTEGSSAFLPCSFN-----ASQRLAIGSVTWFRDEVPGKE---VRNG 51
 DB 21 WGVSSPKNVQGLSGSCLLIFCIPSPADVPVNSGITAI-----WYD--YSGKQVVIHSG 74
 QY 52 TPE-----FRGLAPLASSRFLHDHQ-AELHIRDVRGHDASIVYCRVEV 94
 DB 75 DPKLVDRKFRGRaelMGNM-----DHKVCNLLKDLKPEDSGTYNFRFEI 119

RESULT 14
 US-09-252-150-9
 ; Sequence 59, Application US/09252150A
 ; Patent No. US20020155604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ledbetter, Jeffrey A.
 ; APPLICANT: Hayden Ledbetter, Martha
 ; APPLICANT: Brady, William A.
 ; APPLICANT: Grosmoire, Laura S.
 ; APPLICANT: Law, Che-Jeung
 ; APPLICANT: Dua, Raj
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING
 ; TITLE OF INVENTION: LYMPHOCYTE ACTIVATION
 ; FILE REFERENCE: 9113-0019-999
 ; CURRENT APPLICATION NUMBER: US/09/252,150A
 ; CURRENT FILING DATE: 1999-02-18
 ; EARLIER APPLICATION NUMBER: US 60/075,274
 ; EARLIER FILING DATE: 1998-02-19
 ; EARLIER APPLICATION NUMBER: US 60/108,683
 ; EARLIER FILING DATE: 1998-11-16
 ; NUMBER OF SEQ ID NOS: 80
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 205
 ; TYPE: PRT
 ; ORGANISM: Llama llama
 US-09-252-150-9

Query Match 14.4%; Score 91; DB 9; Length 205;
 Best Local Similarity 28.0%; Pred. No. 0.058;
 Matches 35; Conservative 14; Mismatches 36; Indels 40; Gaps 8;

QY 14 GSSAFLPCSFNASQGR-----LAIGSVTWFRDEVPGKEVNGTPEFRGLAPLASSRFLH 69
 DB 15 GDSRLSLSC---AASGRTEFTNYAMG---WFRQ---APGKE-----PEFVARISRVGSSTFYT 61
 QY 70 D-----HQAELHIRDVRGHDASIVYCRVEVLGLGVGT-----GNGTRL 107
 DB 62 DSVKGRFTISRDNKNTMYLQNSMKAEDTAVYICAAD-SDYGPGRSRSEYDYWGQGTQV 120
 QY 108 VVEKE 112
 DB 121 TVSSSE 125

RESULT 15
 US-09-880-748-1545
 ; Sequence 1545, Application US/09880748

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 23.2361 Seconds
(without alignments)
266.616 Million cell updates/sec

Title: US-10-036-444-4
Perfect score: 632
Sequence: 1 LWYSQPPRTLEGSSAFPLP.....TNGRTLVVEKHPQLGACT 120

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	16.1	267	1	US-08-416-336-2
2	99	15.7	270	3	US-09-082-593-10
3	95.5	15.1	218	4	US-10-026-045-1
4	94	14.9	526	4	US-09-910-174B-9
5	94	14.9	526	4	US-09-620-461-9
6	94	14.9	589	2	US-08-724-394A-1
7	92	14.6	174	3	US-08-804-180C-4
8	89.5	14.2	120	3	US-09-171-945-89
9	89.5	14.2	137	3	US-08-804-180C-2
10	89.5	14.2	391	5	PCT-US95-15696-2
11	88	13.9	226	4	US-09-311-764A-32
12	88	13.9	946	5	PCT-US95-08493-13
13	87.5	13.8	120	4	US-09-472-087-102
14	87.5	13.8	187	1	US-08-067-684-14
15	87.5	13.8	187	1	US-08-008-898-14
16	87.5	13.8	187	2	US-08-459-818-14
17	87.5	13.8	187	2	US-08-889-666-14
18	87.5	13.8	187	2	US-08-485-078-14
19	87.5	13.8	187	2	US-08-725-776-14
20	87.5	13.8	187	2	US-08-488-062-14
21	87.5	13.8	187	3	US-08-228-208A-14
22	87.5	13.8	187	5	PCT-US95-06726-36
23	87.5	13.8	218	4	US-10-026-045-3
24	87.5	13.8	223	3	US-08-208-208A-17
25	87.5	13.8	247	5	PCT-US94-10257A-2
26	87.5	13.8	364	4	US-09-472-087-100
27	87.5	13.8	374	4	US-09-227-595-26

28	87.5	13.8	374	4	US-09-227-595-28	Sequence 28, Appl
29	87.5	13.8	377	4	US-09-227-595-24	Sequence 24, Appl
30	86.5	13.7	218	4	US-10-026-045-2	Sequence 2, Appl
31	86.5	13.7	1297	3	US-09-540-245A-17	Sequence 17, Appl
32	85.5	13.5	120	3	US-09-171-945-81	Sequence 81, Appl
33	85.5	13.5	223	4	US-09-303-040-10	Sequence 10, Appl
34	84	13.3	121	1	US-08-388-672A-20	Sequence 20, Appl
35	84	13.3	126	3	US-09-080-554-20	Sequence 20, Appl
36	84	13.3	360	4	US-09-907-754A-213	Sequence 213, App
37	84	13.3	360	4	US-09-905-125A-213	Sequence 213, App
38	84	13.3	360	4	US-09-902-775A-213	Sequence 213, App
39	83.5	13.2	120	3	US-09-171-945-79	Sequence 79, Appl
40	83.5	13.2	120	3	US-09-171-945-91	Sequence 91, Appl
41	82.5	13.1	124	3	US-08-630-172-4	Sequence 4, Appl
42	82.5	13.1	124	3	US-09-375-419-4	Sequence 4, Appl
43	82.5	13.1	238	4	US-09-227-595-30	Sequence 30, Appl
44	82.5	13.1	238	4	US-09-227-595-32	Sequence 32, Appl
45	82.5	13.1	357	3	US-08-630-172-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-416-336-2
; Sequence 2, Application US/08416336
; Patent No. 5807714
; GENERAL INFORMATION:
; APPLICANT: Ishizaka, Kimishige
; APPLICANT: Ishii, Yasuyuki
; TITLE OF INVENTION: METHOD OF PRODUCTION OF ANTIGEN-SPECIFIC
; TITLE OF INVENTION: GLYCOSYLATION INHIBITING FACTOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,336
FILING DATE: 04-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa H.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07246/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-336-2

Query Match 16.1%; Score 101.5; DB 1; Length 267;
Best Local Similarity 30.2%; Pred. No. 0.00093;
Matches 38; Conservative 16; Mismatches 39; Indels 33; Gaps 8;
QY 2 WYS-----QPPEIRTEGSSAFPLPCSFNASQSGRLAIGSVTWFRDEVVPGKEVRN-----50
DB 18 WYSGDKVQKSSALSQEGTNSALCNPS-----IAATTVQWFLQN--PRGSLMNLFLV 70
QY 51 -GTPBFRGRGLAPLASSRFLHDHQAELHTRDVRGHDASIVYCRVEVLGVGT-----GN 103

Db 71 PGTKE-NGLKSTFNSK---ESYSTLHRAQLEDSTYFCABE-----GGGSNYKLTFGK 122
 Qy 104 GTRLVW 109
 Db 123 GTLLTV 128

RESULT 2
 US-09-082-593-10
 ; Sequence 10, Application US/09082593
 ; Patent No. 6180104
 ; GENERAL INFORMATION:
 ; APPLICANT: DAVIS, MARK M.
 ; APPLICANT: HEDRICK, STEPHEN M.
 ; TITLE OF INVENTION: T CELL RECEPTOR BETA SUBUNIT
 ; FILE REFERENCE: JX1193-195DIV2
 ; CURRENT APPLICATION NUMBER: US/09/082,593
 ; CURRENT FILING DATE: 1998-05-20
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 270
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-082-593-10

Query Match 15.7%; Score 99; DB 3; Length 270;
 Best Local Similarity 29.4%; Pred. No. 0.0018;
 Matches 37; Conservative 17; Mismatches 42; Indels 30; Gaps 8;
 Qy 3 VSOPPE-IRTEGSAFLPCSFNASQGLAIGSVTWFRDEVPVQKEVR-----NGTPE 54
 Db 24 VQSPESILVPEGARTSLNCTFSDASQY-----FWYRQH--SGKAPKALMSIFSNGEKE 77
 Qy 55 FRGLAPLASSRFLHDHQAELHIRDVGRGHDAIYVCRVEVLGVLGVTGNGTRLVVEKEHP 114
 Db 78 -EGRTIHLNKASLH---FSLHSDQPSDSALYLCAVTLVG-----GSGNKLI----- 122
 Qy 115 QLQAGT 120
 Db 123 -FGTGT 127

RESULT 3
 US-10-026-045-1
 ; Sequence 1, Application US/10026045
 ; Patent No. 6573236
 ; GENERAL INFORMATION:
 ; APPLICANT: Genain, Claude P
 ; APPLICANT: Hauser, Stephen L
 ; TITLE OF INVENTION: Inhibiting MOG-Antibody Binding
 ; FILE REFERENCE: UCSF99-020-3
 ; CURRENT APPLICATION NUMBER: US/10/026,045
 ; CURRENT FILING DATE: 2003-01-13
 ; PRIOR APPLICATION NUMBER: US 09/384,036
 ; PRIOR FILING DATE: 1999-08-26
 ; PRIOR APPLICATION NUMBER: us 60/097,953
 ; PRIOR FILING DATE: 1998-08-26
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 218
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 US-10-026-045-1

Query Match 15.1%; Score 95.5; DB 4; Length 218;
 Best Local Similarity 36.2%; Pred. No. 0.0035;
 Matches 34; Conservative 8; Mismatches 33; Indels 19; Gaps 6;
 Qy 9 IRTLEGSAFLPCSFNASQGLAIG-SVTWFRDEVPVQKEV-----RNG-----TPFPR 56

Db 12 IRALVGDABEALPC--RISPGKNATGMEVGWYRS---PFSRVVHLRYRNGKQDAEQAPEYR 66
 Qy 57 GRPLAPLASSRFLHDHQAELHIRDVGRGHDAIYVC 90
 Db 67 GRTELLKES--IGEGKVALRIQNVRFSDGGYTC 98

RESULT 4
 US-09-910-174B-9
 ; Sequence 9, Application US/09910174B
 ; Patent No. 6630575
 ; GENERAL INFORMATION:
 ; APPLICANT: Coyle, Anthony J.
 ; APPLICANT: Fraser, Christopher C.
 ; APPLICANT: Manning, Stephen
 ; TITLE OF INVENTION: B7-H2 Molecules, No. 6630575el Members of the B7
 ; TITLE OF INVENTION: Family and Uses Thereof
 ; FILE REFERENCE: 35800/236924
 ; CURRENT APPLICATION NUMBER: US/09/910,174B
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 09/620,461
 ; PRIOR FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 526
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-910-174B-9

Query Match 14.9%; Score 94; DB 4; Length 526;
 Best Local Similarity 39.6%; Pred. No. 0.017;
 Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;
 Qy 6 PPE-IRTEGSAFLPC--SFNASQGLAIGSVTWFRDEVPVQKEV-RNG-----TPE 54
 Db 34 PPEPILAVVGEDAEALFCRLSPNASEHL---ELRFRKKVSPAVLVHRDGRQEAQMP 90
 Qy 55 FRGLAPLASSRFLHDHQAELHIRDVGRGHDAIYVC 90
 Db 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDDGEYTC 124

RESULT 5
 US-09-620-461-9
 ; Sequence 9, Application US/09620461
 ; Patent No. 6635750
 ; GENERAL INFORMATION:
 ; APPLICANT: Coyle, Anthony J.
 ; APPLICANT: Fraser, Christopher C.
 ; APPLICANT: Manning, Stephen
 ; TITLE OF INVENTION: B7-H2 Molecules, No. 6635750el Members of the B7
 ; TITLE OF INVENTION: Family and Uses Thereof
 ; FILE REFERENCE: 5800-149
 ; CURRENT APPLICATION NUMBER: US/09/620,461
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 526
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-620-461-9

Query Match 14.9%; Score 94; DB 4; Length 526;
 Best Local Similarity 39.6%; Pred. No. 0.017;
 Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;
 Qy 6 PPE-IRTEGSAFLPC--SFNASQGLAIGSVTWFRDEVPVQKEV-RNG-----TPE 54
 Db 34 PPEPILAVVGEDAEALFCRLSPNASEHL---ELRFRKKVSPAVLVHRDGRQEAQMP 90
 Qy 55 FRGLAPLASSRFLHDHQAELHIRDVGRGHDAIYVC 90

Db 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDGGEYTC 124

RESULT 6
US-08-724-394A-1
; Sequence 1, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kromal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wollf, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724.394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 589 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..589
; OTHER INFORMATION: /note= "BT"
US-08-724-394A-1

Query Match 14.9%; Score 94; DB 2; Length 589;
Best Local Similarity 39.6%; Pred. No. 0.019;
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;
QY 6 PPE-IRTEGSSAFPC--SFNASQRLAIGSVTWFRDEVVPGKEV-RNG-----TPE 54
Db 37 PPEPLAVVGDEALPCRLSPNASAEHL---ELNFRKKVSPAVLVHRDGRGEAEQWPE 93

QY 55 FGRLAPLASSRFLHDHQAELHIRDVGRHDASIYVC 90
Db 94 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDGGEYTC 127

RESULT 7
US-08-804-180C-4
; Sequence 4, Application US/08804180C
; Patent No. 6107056
; GENERAL INFORMATION:
; APPLICANT: Emery, Stephen
; APPLICANT: Copley, Clive Graham
; APPLICANT: Edge, Michael Derek
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

Query Match 14.6%; Score 92; DB 3; Length 174;
Best Local Similarity 29.5%; Pred. No. 0.0067;
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;
QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58
Db 40 VTQPSVVLASSHGVASFPCVYASHNTDEVRTVLRQT--NDQVT---EVCAITFTVKNT 94

QY 59 LA----PLASSRFLHDHQAELHIRDVGRHDASIYVCVEVL---GLGVGTGNGTGL-VVE 110
Db 95 LGFLDDPFCGTF--NESRVNLTIOGLRAADTGLYFCKVELMYPFPYFGMGNGTQIYVIA 153

QY 111 KE 112
Db 154 KE 155

RESULT 8
US-09-171-945-89
; Sequence 89, Application US/09171945
; Patent No. 6277599
; GENERAL INFORMATION:
; APPLICANT: Emery, Stephen
; APPLICANT: Copley, Clive Graham
; APPLICANT: Edge, Michael Derek
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

Query Match 14.6%; Score 92; DB 3; Length 174;
Best Local Similarity 29.5%; Pred. No. 0.0067;
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;
QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58
Db 40 VTQPSVVLASSHGVASFPCVYASHNTDEVRTVLRQT--NDQVT---EVCAITFTVKNT 94

QY 59 LA----PLASSRFLHDHQAELHIRDVGRHDASIYVCVEVL---GLGVGTGNGTGL-VVE 110
Db 95 LGFLDDPFCGTF--NESRVNLTIOGLRAADTGLYFCKVELMYPFPYFGMGNGTQIYVIA 153

APPLICANT: Martin K. Oaks
TITLE OF INVENTION: SCTLA-4 and Its Soluble Products
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thomas M. Wozny
STREET: 100 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.50 inch Disk
COMPUTER: IBM
OPERATING SYSTEM: DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.180C
FILING DATE: February 20, 1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Thomas M. Wozny
REGISTRATION NUMBER: 28,922
REFERENCE/DOCKET NUMBER: 3284-00003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 271-7590
TELEFAX: (414) 271-5770

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 174
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: Polypeptide
HYPOTHETICAL: no
FRAGMENT TYPE: Internal fragment
ORIGINAL SOURCE:
ORGANISM: Rattus No. 6107056vegicus
STRAIN: ACI
DEVELOPMENTAL STAGE: Adult
CELL TYPE: Splenocyte
FEATURE:
NAME/KEY: Rat SCTLA-4
IDENTIFICATION METHOD: Found by experiment
OTHER INFORMATION: B7 binding protein
US-08-804-180C-4

Query Match 14.6%; Score 92; DB 3; Length 174;
Best Local Similarity 29.5%; Pred. No. 0.0067;
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;
QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58
Db 40 VTQPSVVLASSHGVASFPCVYASHNTDEVRTVLRQT--NDQVT---EVCAITFTVKNT 94

QY 59 LA----PLASSRFLHDHQAELHIRDVGRHDASIYVCVEVL---GLGVGTGNGTGL-VVE 110
Db 95 LGFLDDPFCGTF--NESRVNLTIOGLRAADTGLYFCKVELMYPFPYFGMGNGTQIYVIA 153

QY 111 KE 112
Db 154 KE 155

RESULT 8
US-09-171-945-89
; Sequence 89, Application US/09171945
; Patent No. 6277599
; GENERAL INFORMATION:
; APPLICANT: Emery, Stephen
; APPLICANT: Copley, Clive Graham
; APPLICANT: Edge, Michael Derek
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

Query Match 14.6%; Score 92; DB 3; Length 174;
Best Local Similarity 29.5%; Pred. No. 0.0067;
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;
QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58
Db 40 VTQPSVVLASSHGVASFPCVYASHNTDEVRTVLRQT--NDQVT---EVCAITFTVKNT 94

QY 59 LA----PLASSRFLHDHQAELHIRDVGRHDASIYVCVEVL---GLGVGTGNGTGL-VVE 110
Db 95 LGFLDDPFCGTF--NESRVNLTIOGLRAADTGLYFCKVELMYPFPYFGMGNGTQIYVIA 153

QY 111 KE 112
Db 154 KE 155

RESULT 8
US-09-171-945-89
; Sequence 89, Application US/09171945
; Patent No. 6277599
; GENERAL INFORMATION:
; APPLICANT: Emery, Stephen
; APPLICANT: Copley, Clive Graham
; APPLICANT: Edge, Michael Derek
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said
; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

Query Match 14.6%; Score 92; DB 3; Length 174;
Best Local Similarity 29.5%; Pred. No. 0.0067;
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;
QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58
Db 40 VTQPSVVLASSHGVASFPCVYASHNTDEVRTVLRQT--NDQVT---EVCAITFTVKNT 94

FILE REFERENCE: Monoclonal Antibody to CEA
 CURRENT APPLICATION NUMBER: US/09/171,945
 CURRENT FILING DATE: 1998-10-29
 PRIOR APPLICATION NUMBER: GB9703103.3
 PRIOR FILING DATE: 1997-02-14
 PRIOR APPLICATION NUMBER: GB9609405.7
 PRIOR FILING DATE: 1996-05-04
 PRIOR APPLICATION NUMBER: PCT/GB97/01165
 PRIOR FILING DATE: 1997-04-29
 NUMBER OF SEQ ID NOS: 131
 SOFTWARE: Patent in Ver. 2.1
 SEQ ID NO 89
 LENGTH: 120
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: humanized
 US-09-171-945-89

Query Match 14.2%; Score 89.5; DB 3; Length 120;
 Best Local Similarity 27.8%; Pred. No. 0.008;
 Matches 32; Conservative 10; Mismatches 34; Indels 39; Gaps 6;

QY 2 WVSOPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVPQKEVRNG---TPFRG 57
 DB 36 WVRQPP-----GR-GLGWIGWIDPE-----NGDTEYAPKFRG 65
 QY 58 RLAPLASSRFLHDOAEHLHIDVRGHDASIVYCVREVLGLGVGT---GNGTRLVW 109
 DB 67 RATMLADS---SKNQASLELSVTAADTAIVYCHVLIYAGYALMDYWGQGLTVV 118

RESULT 9

US-08-804-180C-2
 Sequence 2, Application US/08804180C
 Patent No. 6107056

GENERAL INFORMATION:
 APPLICANT: Martin K. Oaks
 TITLE OF INVENTION: SCTL4-4 and Its Soluble Products
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Thomas M. Wozny
 STREET: 100 East Wisconsin Avenue
 CITY: Milwaukee
 STATE: Wisconsin
 COUNTRY: USA
 ZIP: 53202

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.50 inch Disk
 COMPUTER: IBM
 OPERATING SYSTEM: DOS
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/804,180C
 FILING DATE: February 20, 1997
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Thomas M. Wozny
 REGISTRATION NUMBER: 28,922
 REFERENCE/DOCKET NUMBER: 3284-00003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (414) 271-7590
 TELEFAX: (414) 271-5770

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 137
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 DESCRIPTION: no
 HYPOTHETICAL: no
 FRAGMENT TYPE: Mature Polypeptide
 ORIGINAL SOURCE:

ORGANISM: Homo Sapien
 DEVELOPMENTAL STAGE: Adult
 TISSUE TYPE: Lymphnode
 FEATURE:
 NAME/KEY: Human SCTL4-4
 IDENTIFICATION METHOD: Found by experiment
 OTHER INFORMATION: Asn 76 and Asn 108 represent N-linked glycosylation; B7 bind
 US-08-804-180C-2

Query Match 14.2%; Score 89.5; DB 3; Length 137;
 Best Local Similarity 25.2%; Pred. No. 0.0095;
 Matches 33; Conservative 18; Mismatches 59; Indels 21; Gaps 4;

QY 3 VSQPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVPGKEVRNGTPEFRGLAPL 62
 DB 3 VAQPAVLASSRGIASFVCEY-ASPGKATEVRVTVLRQADSQVTEVCAATMMGNELT-- 59
 QY 63 ASSRFLHD-----HQAELHIDVRGHDASIVYCVREVL-----GLGVGTGNGTRLV 108
 DB 60 ----FLDSDICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPPYLGIENGTOIYVI 115
 QY 109 VEKEHPQLGAG 119
 DB 116 AXEKPSYNRG 126

RESULT 10

PCT-US95-15696-2
 Sequence 2, Application PC/TUS9515696
 GENERAL INFORMATION:

APPLICANT: President and Fellows of Harvard College
 TITLE OF INVENTION: SINGLE CHAIN T-CELL RECEPTOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kevin M. Farrell, P.C.
 STREET: P.O. Box 999
 CITY: York Harbor
 STATE: Maine
 COUNTRY: USA
 ZIP: 03911

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/15696
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/349,915
 FILING DATE: 06-DEC-1994

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Farrell, Kevin M.
 REGISTRATION NUMBER: 35,505
 REFERENCE/DOCKET NUMBER: HU-9404 WO
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 207-363-0558
 TELEFAX: 207-363-0528
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 391 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-15696-2

Query Match 14.2%; Score 89.5; DB 5; Length 391;
 Best Local Similarity 27.6%; Pred. No. 0.037;
 Matches 34; Conservative 17; Mismatches 49; Indels 23; Gaps 7;

QY 6 PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVPGKEV-----RNGTPEFRGR 58

Db 60 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLIGNGTQIYVI 115
QY 111 KEHP 114
Db 116 DPEP 119

RESULT 14

US-08-067-684-14
; Sequence 14, Application US/08067684
; Patent No. 5434131
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Suite 900
; CITY: Pasadena
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 91101

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/067,684
; FILING DATE: 26-MAY-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Adriano, Sarah B.
; REGISTRATION NUMBER: 34,470
; REFERENCE/DOCKET NUMBER: 7848-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310/312-9900
; TELEFAX: 310/479-8340
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-067-684-14

Query Match 13.8%; Score 87.5; DB 1; Length 187;
Best Local Similarity 27.4%; Pred. No. 0.024; Mismatches 15; Indels 19; Gaps 4;
Matches 34; Conservative 15; Mismatches 56; Indels 19; Gaps 4;
QY 3 VSQPEIRTLGSSAPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGR LAPL 62
Db 4 VAQPAVTLASSRGIA SFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNELT-- 60
QY 63 ASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL---GLGVGTGNGTRELVE 110
Db 61 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLIGNGTQIYVI 116
QY 111 KEHP 114
Db 117 DPEP 120

RESULT 15

US-08-008-898-14
; Sequence 14, Application US/08008898
; Patent No. 5770197
; GENERAL INFORMATION:

; APPLICANT: Linsley, Peter S.
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Damle, Nitin K.
; APPLICANT: Brady, William
; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 201 South Lake Avenue, Suite 800
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,898
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/723,617
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandel, Saralynn
; REGISTRATION NUMBER: 31,853
; REFERENCE/DOCKET NUMBER: 7848
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-008-898-14

Query Match 13.8%; Score 87.5; DB 1; Length 187;
Best Local Similarity 27.4%; Pred. No. 0.024; Mismatches 15; Indels 19; Gaps 4;
Matches 34; Conservative 15; Mismatches 56; Indels 19; Gaps 4;
QY 3 VSQPEIRTLGSSAPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGR LAPL 62
Db 4 VAQPAVTLASSRGIA SFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNELT-- 60
QY 63 ASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL---GLGVGTGNGTRELVE 110
Db 61 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLIGNGTQIYVI 116
QY 111 KEHP 114
Db 117 DPEP 120
Search completed: February 26, 2004, 12:16:15
Job time : 24.2361 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 2.87268 Seconds
(without alignments)
636.214 Million cell updates/sec

Title: US-10-036-444-5
Perfect score: 88
Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	53.4	835	T06590	probable beta-gala
2	46	52.3	724	T04340	beta-galactosidase
3	46	52.3	757	T02561	probable cellulose
4	44.5	50.6	479	T31293	aldehyde dehydroge
5	44	50.0	184	C90157	hypothetical prote
6	43	48.9	466	AH1800	transmembrane effl
7	43	48.9	581	S03540	gene frizzled prot
8	42	47.7	217	T69512	conserved hypothet
9	42	47.7	271	T88918	hypothetical prote
10	42	47.7	388	B83228	probable MFS trans
11	42	47.7	466	A11426	transmembrane effl
12	42	47.7	560	S46096	probable membrane
13	42	47.7	658	S74246	sulfate transport
14	41	46.6	310	F71027	hypothetical prote
15	41	46.6	333	A70602	hypothetical prote
16	41	46.6	356	A70332	hypothetical prote
17	41	46.6	364	B83787	hypothetical prote
18	41	46.6	451	E90171	inorganic phosphat
19	41	46.6	498	T11901	hypothetical prote
20	40	45.5	135	B83466	conserved hypothet
21	40	45.5	154	T69477	NADH2 dehydrogenas
22	40	45.5	233	C95987	probable two-compo
23	40	45.5	245	I51323	proteolipid protei
24	40	45.5	288	T37709	hypothetical prote
25	40	45.5	303	T42703	hypothetical prote
26	40	45.5	328	C87673	4-hydroxybenzoate
27	40	45.5	342	AG2092	hypothetical prote
28	40	45.5	351	G71983	hypothetical prote
29	40	45.5	351	E64524	NADH2 dehydrogenas

30 40 45.5 382 2 H86930 probable secreted
31 40 45.5 401 2 F75037 hexuronate transpo
32 40 45.5 412 2 F97196 probable permease,
33 40 45.5 423 2 G71850 proline/betaine tr
34 40 45.5 424 2 F97870 probable betaine tr
35 40 45.5 452 2 T45448 probable serine pr
36 40 45.5 507 2 F90540 hypothetical prote
37 40 45.5 514 2 S46733 hypothetical prote
38 40 45.5 572 2 T13740 probable hormone r
39 40 45.5 606 2 AC2425 ATP-binding protei
40 40 45.5 610 2 F91025 probable transport
41 40 45.5 610 2 A85870 probable transport
42 40 45.5 610 2 B65001 probable transport
43 40 45.5 1040 2 D81379 transmembrane effl
44 39.5 44.9 396 2 AD3012 conserved hypothet
45 39.5 44.9 492 2 C98272 hypothetical prote

ALIGNMENTS

RESULT 1

T06590
Probable beta-galactosidase (EC 3.2.1.23) - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000
C:Accession: T06590
R:Carey, A.T.; Holt, K.; Picard, S.; Wilde, R.; Tucker, G.A.; Bird, C.R.; Schuch, W.; S
Plant Physiol. 108, 1099-1107, 1995
A:Title: Tomato exo-(1-4)-beta-D-galactanase: isolation, changes during ripening in nor
A:Reference number: Z15780; MUID:95357407; PMID:7630937
A:Accession: T06590
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-835 <CAF>
A:Cross-references: EMBL:X83854; NID:G971484; PIDN:CAA58734.1; PID:G971485
A:Experimental source: cultivar Ailsa Craig; pericarp
C:Superfamily: beta-galactosidase bga
C:Keywords: Glycosidase; hydrolase

Query Match 53.4%; Score 47; DB 2; Length 835;
Best Local Similarity 62.5%; Pred. No. 8;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19
DB 527 LRAGVKNLSLSIAGV 542

RESULT 2

T04340
beta-galactosidase (EC 3.2.1.23) II precursor - tomato
C:Species: Lycopersicon esculentum (tomato)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000
C:Accession: T04340
R:Smith, D.L.; Starrett, D.A.; Gross, K.C.
Plant Physiol. 117, 417-423, 1998
A:Title: A gene coding for tomato fruit beta-galactosidase II is expressed during fruit
A:Reference number: Z15296; MUID:98289087; PMID:9625694
A:Accession: T04340
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-724 <SMT>
A:Cross-references: EMBL:AF020390; NID:G3299895; PIDN:AAC5984.1; PID:G3299896
A:Experimental source: strain Rutgers; tissue-type tomato fruit
C:Genetics:
A:Gene: Bgal4
C:Superfamily: beta-galactosidase bga
C:Keywords: Glycosidase; hydrolase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-724/Product: beta-galactosidase II #status predicted <MAT>

Query Match 52.3%; Score 46; DB 2; Length 724;

Best Local Similarity 62.5%; Pred. No. 10; Mismatches 2; Indels 0; Gaps 0;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19
||||| :||||| :||
Db 529 LRAGINKISLSSVSVG 544
||||| :||||| :||

RESULT 3
T02561
probable cellulose synthase A2g32620 [imported] - Arabidopsis thaliana
N;Alternate names: cellulose synthase homolog T26B15.18
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02561; D84735
P;Roundley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.
A;Reference number: Z14678
A;Accession: T02561
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-757 <RCU>
A;Cross-references: EMBL:AC004681; NID:g3298532; PID:g3298550
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niekman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84735
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-757 <STC>
A;Cross-references: GB:AE002093; NID:g3298550; PIDN:AAAC25944.1; GSPDB:GN00139
C;Genetics:
A;Gene: Atg32620; T26B15.18
A;Map position: 2
A;Introns: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 566/3

Query Match 52.3%; Score 46; DB 2; Length 757;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRAGFYAVSFLSVAVG 19
|:|||||:|
Db 740 LKAGFLAVSFVFSVG 756
|:|||||:|

RESULT 4
T31293
aldehyde dehydrogenase homolog - Sphingomonas aromaticivorans plasmid pNL1
C;Species: Sphingomonas aromaticivorans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C;Accession: T31293
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G
submitted to the EMBL Data Library, July 1998
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati
A;Reference number: Z20992
A;Accession: T31293
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-479 <ROW>
A;Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378434; PIDN:AAD04017.1
C;Genetics:
A;Genome: plasmid pNL1
A;Note: nanF
C;Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 50.6%; Score 44.5; DB 2; Length 479;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 3 LLRAGF---YAVSFLSVAVG 19
:|||||:|||||:|
Db 95 IAKAGFETFAVSLRAAIG 114
:|||||:|||||:|

RESULT 5
C90157
hypothetical protein SS00169 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: C90157
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90157
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-184 <KUR>
A;Cross-references: GB:AE006641; NID:g13813300; PIDN:AAK40514.1; GSPDB:GN00155
C;Genetics:
A;Gene: SS00169

Query Match 50.0%; Score 44; DB 2; Length 184;
Best Local Similarity 75.0%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LLRAGFYAVSFL 14
|:|||||:|
Db 76 LISAGLYAVSFL 87
|:|||||:|

RESULT 6
AH1800
transmembrane efflux protein homolog lin2951 [imported] - Listeria innocua (strain Clif
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 06-Jan-2003
C;Accession: AH1800
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlianc
A;Title: Comparative Genomics of Listeria species
A;Reference number: AS1077; MUID:21537279; PMID:11679669
A;Accession: AH1800
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC98176.1; PID:g16415492; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2951
C;Superfamily: multidrug-efflux transporter

Query Match 48.9%; Score 43; DB 2; Length 466;
Best Local Similarity 58.8%; Pred. No. 21;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLRAGFYAVSFLSVAV 18
|||:|||||:|
Db 434 LLTNVGFVCVSLISVAV 450
|||:|||||:|

RESULT 7
S03540
gene frizzled protein precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 02-Mar-2001

C;Accession: S03540; S15709; S15709
 R;Vinson, C.R.; Conover, S.; Adler, P.N.
 Nature 338, 263-264, 1989
 A;Title: A Drosophila tissue polarity locus encodes a protein containing seven potential
 A;Reference number: S03540; MUID:89159415; PMID:2493583
 A;Accession: S03540
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-581 <VIN>
 R;Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
 Genetics 126, 401-416, 1990
 A;Title: Molecular structure of frizzled, a Drosophila tissue polarity gene.
 A;Reference number: S15708; MUID:91060073; PMID:2174014
 A;Accession: S15708
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-581 <ADL>
 A;Cross-references: EMBL:X54648; NID:G7973; PIDN:CAA38460.1; PID:G804979
 A;Accession: S15709
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-405, 'WY', 408, 'WQHTIN' <AD2>
 A;Cross-references: EMBL:X54648; NID:G7973; PIDN:CAA38461.1; PID:G804980
 C;Genetics:
 A;Gene: FlyBase: fz
 A;Cross-references: FlyBase:FBgn0001085
 A;Introns: 224/3; 264/3; 329/3; 405/1
 C;Superfamily: fruit fly frizzled protein
 C;Keywords: alternative splicing; transmembrane protein
 F;1-26/Domain: signal sequence #status predicted <SIG>
 F;27-581/Product: gene frizzled protein #status predicted <WAT>
 Query Match 48.9%; Score 43; DB 2; Length 581;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 2 LLLRAGFYAVSFLSVAVG 19
 DB 468 LMLRIGFFSGLFILPAVG 485
 RESULT 8
 F69512
 conserved hypothetical protein AF2102 - Archaeoglobus fulgidus
 C;Species: Archaeoglobus fulgidus
 C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.P.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A;Reference number: A69250; MUID:98049343; PMID:9389475
 A;Accession: F69512
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-217 <KLB>
 A;Cross-references: GB:A8000959; GB:A8000782; NID:G2689281; PIDN:AB89144.1; PID:G264842
 Query Match 47.7%; Score 42; DB 2; Length 217;
 Best Local Similarity 76.9%; Pred. No. 14;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 LLLRAGFYAVSFL 14
 DB 198 LLLRAGLLAVSFL 210
 RESULT 9
 B86918
 hypothetical protein glpQ [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C;Accession: B86918
 R;Coile, S.T.; Eigemeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; F
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, I.; Fraser, A.; Hamlin, N.; Holroyd
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
 A;Title: Massive gene decay in the leprosy bacillus.
 A;Reference number: A86909; MUID:21128732; PMID:11234002
 A;Accession: B86918
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-271 <STO>
 A;Cross-references: GB:AL450380; NID:G13092466; PIDN:CAC29582.1; GSPDB:GN00147
 C;Genetics:
 A;Gene: glpQ
 Query Match 47.7%; Score 42; DB 2; Length 271;
 Best Local Similarity 57.9%; Pred. No. 18;
 Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 1 VLLLRAGFYAVSFLSVAVG 19
 DB 184 VLLGRAGEYLTSSAATAVG 202
 RESULT 10
 B83228
 probable MFS transporter PA3336 [imported] - Pseudomonas aeruginosa (strain PA01)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; E
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lin
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: B83228
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-388 <STO>
 A;Cross-references: GB:AE004756; GB:AE004091; NID:G9949466; PIDN:AAG06724.1; GSPDB:GN0
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA3336
 C;Superfamily: Streptomycetes lividans chloramphenicol resistance protein
 Query Match 47.7%; Score 42; DB 2; Length 388;
 Best Local Similarity 53.3%; Pred. No. 26;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 5 RAGFYAVSFLSVAVG 19
 DB 157 RASFFAVAVLVG 171
 RESULT 11
 AL1426
 transmembrane efflux protein homolog lmo2818 [imported] - Listeria monocytogenes (stra
 C;Species: Listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 06-Jan-2003
 C;Accession: AL1426
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, I
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; N
 ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant
 A;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AL1426

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD01031.1; PID:G16412318; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmc2818
C;Superfamily: multidrug-efflux transporter

Query Match 47.7%; Score 42; DB 2; Length 466;
Best Local Similarity 52.9%; Pred. No. 31;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAV 18
DB 434 LTTNAGFCVSLISVAI 450

RESULT 12
S46096
probable membrane protein YBR220c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR1510
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 19-Apr-2002
C;Accession: S46096
R;Dubois, E.; El Bakkoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.; submitted to the Protein Sequence Database, August 1994
A;Reference number: S45782
A;Accession: S46096
A;Molecule type: DNA
A;Residues: 1-560 <DUB>
A;Cross-references: EMBL:Z36088; NID:G536609; PID:G536611; GSPDB:GN00002; MIPS:YBR220c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YBR220c
A;Cross-references: SGD:S0000424
A;Map position: 2R
C;Keywords: transmembrane protein
F;18-37/Domain: transmembrane #status predicted <TM01>
F;188-106/Domain: transmembrane #status predicted <TM02>
F;139-155/Domain: transmembrane #status predicted <TM03>
F;177-193/Domain: transmembrane #status predicted <TM04>
F;218-238/Domain: transmembrane #status predicted <TM05>
F;330-347/Domain: transmembrane #status predicted <TM06>
F;391-407/Domain: transmembrane #status predicted <TM07>
F;421-437/Domain: transmembrane #status predicted <TM08>
F;475-491/Domain: transmembrane #status predicted <TM09>
F;521-544/Domain: transmembrane #status predicted <TM10>

Query Match 47.7%; Score 42; DB 2; Length 560;
Best Local Similarity 36.8%; Pred. No. 37;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLLRAGFYAVSFLSVAV 19
DB 516 VTLRDGYITNLICIVG 534

RESULT 13
S74246
sulfate transport protein - Arabidopsis thaliana
N;Alternate names: sulfate transporter
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Oct-1996 #sequence_revision 14-Nov-1997 #text_change 20-Jun-2000
C;Accession: S74246
R;Takahashi, H.; Sasakura, N.; Noji, M.; Saito, K.
FEBS Lett. 392, 95-99, 1996
A;Title: Isolation and characterization of a cDNA encoding a sulfate transporter from Arabidopsis thaliana
A;Reference number: S74246; MUID:96368029; PMID:8772182
A;Accession: S74246
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-658 <TAK>

A;Cross-references: EMBL:D85416; NID:G1498119; PID:G1498120
C;Superfamily: sulfate transport protein
C;Keywords: sulfate transporter; transmembrane protein
F;92-109/Domain: transmembrane #status predicted <TM1>
F;119-134/Domain: transmembrane #status predicted <TM2>
F;140-156/Domain: transmembrane #status predicted <TM3>
F;172-190/Domain: transmembrane #status predicted <TM4>
F;194-217/Domain: transmembrane #status predicted <TM5>
F;253-271/Domain: transmembrane #status predicted <TM6>
F;278-293/Domain: transmembrane #status predicted <TM7>
F;338-353/Domain: transmembrane #status predicted <TM8>
F;371-388/Domain: transmembrane #status predicted <TM9>
F;408-430/Domain: transmembrane #status predicted <TM10>
F;435-457/Domain: transmembrane #status predicted <TM11>
F;464-489/Domain: transmembrane #status predicted <TM12>

Query Match 47.7%; Score 42; DB 2; Length 658;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAV 19
DB 465 LVLLNAFFGVLFASVEIG 482

RESULT 14
F71027
hypothetical protein PH1514 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C;Accession: F71027
R;Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: F71027
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-310 <KAW>
A;Cross-references: GB:AF000006; NID:G3236133; PIDN:BAA30622.1; PID:G3257939
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1514

Query Match 46.6%; Score 41; DB 2; Length 310;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 FYAVSFLSVAV 19
DB 34 YYLAFSLGVG 45

RESULT 15
A70602
hypothetical protein Rv0998 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: A70602
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70602
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-333 <COL>
A;Cross-references: GB:Z94752; GB:AL123456; NID:G3261731; PIDN:CAB08156.1; PID:el299775

A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: RV0998

Query Match 46.6%; Score 41; DB 2; Length 333;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LLLRAGFVAVSFLSVAVG 19
:||| | |||| :|
Db 46 VLLRQGEPAVSFLISSG 63

Search completed: February 26, 2004, 12:14:53
Job time : 5.87268 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 1.61273 Seconds
(without alignments)
613.452 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 86

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	53.4	835	1 BGAL_LYCES	P48980 lycopersico
2	44	50.0	381	1 O46A DROME	P81919 drosophila
3	43.5	49.4	355	1 S35D HUMAN	Q9ntn3 homo sapien
4	43	48.9	581	1 FRIZ DROME	P18537 drosophila
5	42	47.7	127	1 PR69 YEAST	P38317 saccharomyc
6	42	47.7	217	1 YL02 ARCFU	C28178 archaeoglob
7	42	47.7	412	1 CCA_EUCAP	Q8ka53 buchnera ap
8	42	47.7	560	1 YB70 YEAST	P38318 saccharomyc
9	42	47.7	658	1 Y222 ARATH	P52946 arabidopsis
10	41	46.6	356	1 Y359 AQUAE	O66685 aquifex aeo
11	41	46.6	498	1 YOH2 CAEEL	Q09502 caenorhabdi
12	40	45.5	123	1 PHF2 RHIME	Q52955 rhizobium m
13	40	45.5	154	1 Y123 ARCFU	C28452 archaeoglob
14	40	45.5	220	1 Y402 PASMU	Q3cmms pasteurilla
15	40	45.5	245	1 DNA_SQUAC	P36983 squallus aca
16	40	45.5	288	1 YDO9 SCHPO	O13729 schizosacch
17	40	45.5	434	1 BC13 MOUSE	P59017 mus musculu
18	40	45.5	485	1 BC13 HUMAN	Q9bxx5 homo sapien
19	40	45.5	514	1 YHK8 YEAST	P38776 saccharomyc
20	40	45.5	610	1 YFBS_ECOLI	P77741 escherichia
21	40	45.5	1107	1 ALA2 ARATH	P38205 arabidopsis
22	39	44.3	68	1 ATPJ HUMAN	P56385 homo sapien
23	39	44.3	70	1 ATPJ_PIG	Q9myt8 sus scrofa
24	39	44.3	397	1 GS9C DROME	Q9w1u5 drosophila
25	39	44.3	634	1 ST35 ARATH	Q941w6 arabidopsis
26	39	44.3	1221	1 YMP3 CAEEL	Q10947 caenorhabdi
27	38	43.2	108	1 NTRB_METBA	Q3v8k3 methanosarc
28	38	43.2	199	1 NUGM_MARPO	P26850 cyanidium c
29	38	43.2	201	1 NUGM_CYACA	P48950 cyanidium c
30	38	43.2	207	1 NUGM_PROWI	Q37626 proteotheca
31	38	43.2	225	1 TRMA_NEICO	P55134 reisserchia
32	38	43.2	232	1 CLB3_SHEEP	P80943 ovine aries
33	38	43.2	236	1 BIR8_GORGO	Q95m71 gorilla gor

34 38 43.2 236 1 BIR8 HUMAN Q96p09 homo sapien
35 38 43.2 236 1 BIR8_PANTR Q95m72 pan troglod
36 38 43.2 258 1 MYPR_ONCMY P79826 oncorhynchu
37 38 43.2 324 1 YR53_CAEEL Q09322 caenorhabdi
38 38 43.2 329 1 Y188_MYCGE P47434 mycoplasma
39 38 43.2 362 1 TRMA_NEIMA Q9jt82 neisseria m
40 38 43.2 403 1 ARGD_VIBVO P59323 vibrio vuln
41 38 43.2 419 1 YHAP_BAGSU Q07523 bacillus su
42 38 43.2 434 1 YRKA_BAGSU P54428 bacillus su
43 38 43.2 511 1 XASA_ECO57 P58229 escherichia
44 38 43.2 511 1 XASA_ECOLI P39183 escherichia
45 38 43.2 513 1 NDII_YEAST P32340 saccharomyc

ALIGNMENTS

RESULT 1

ID BGAL_LYCES STANDARD; PRT; 835 AA.
 AC P48980;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-galactosidase) (Exo-(1-->4)-beta-D-galactanase).
 DE Lycopersicon esculentum (tomato)...
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=CV. Alisa Craig; TISSUE=Pericarp;
 RX MEDLINE=95357407; PubMed=7630937;
 RA Carey A.T., Holt K., Picard S., Wilde R., Tucker G.A., Bird C.R.,
 RT "Tomato exo-(1-->4)-beta-D-galactanase. Isolation, changes during
 RT ripening in normal and mutant tomato fruit, and characterization of a
 RT related cDNA clone".
 RL Plant Physiol. 108:1099-1107(1995).
 CC -!- FUNCTION: Involved in cell wall degradation. Degrades
 CC polysaccharides containing beta-(1-->4)-linked galactans, acting
 CC as an exo-(1-->4)-beta-D-galactanase.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC galactose residues in beta-D-galactosides.
 CC -!- MISCELLANEOUS: Has a pH optimum of 4.5.
 CC -!- SIMILARITY: Belongs to family 35 of glycosyl hydrolases.
 CC -!- SIMILARITY: Contains 1 SUEL-type lectin domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X83854; CAA58734.1; -.
 CC PIR; T06590; T06590.
 CC InterPro; IPR008979; Gal_bind like.
 CC InterPro; IPR000922; Gal_lectin.
 CC InterPro; IPR001944; Glyco_hydro_35.
 CC Pfam; PF02140; Gal_lectin; 1.
 CC Pfam; PF01301; Glyco_hydro_35; 1.
 CC PRINTS; PR00742; GLHYDLASE35.
 CC ProDom; PD005612; Gal_lectin; 1.
 CC PROSITE; PS01182; GLYCOSYL_HYDROL_F35; 1.
 CC PROSITE; PS02028; SUEL_LECTIN; 1.
 CC HydroLase; Glycosidase; Signal.
 CC SIGNAL 1
 CC FT CHAIN 23 835 BETA-GALACTOSIDASE.
 CC FT DOMAIN 749 835 SUEL-TYPE LECTIN.

FT ACT SITE 180 180 PROTON DONOR (POTENTIAL).
 FT ACT_SITE 249 249 NUCLEOPHILE (POTENTIAL).
 SQ SEQUENCE 835 AA; 93336 MW; 94C9685F95C4A46 CRC64;
 Query Match 53.4%; Score 47; DB 1; Length 835;
 Best Local Similarity 62.5%; Pred. No. 4.7;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 4 LRAGFYAVSFSLVAVG 19
 Dd 527 LRAGWKISLSTIAGV 542
 RESULT 2
 O46A_DROME
 ID O46A_DROME STANDARD; PRT; 381 AA.
 AC P81919; Q9U6X9; Q9U5H3.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Odorant receptor 46a, isoform A.
 GN OR46A OR OR46B OR OR46F OR OR46F.1 OR OR46F.2 OR DOR46F OR DOR46F.1
 GN OR DOR19 OR AN8 OR AN9 OR CG17848/CG17849.
 OS Drosophila melanogaster (Fruit fly)
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RX SEQUENCE FROM N.A.
 EX MEDLINE=93189723; PubMed=10458908;
 RA Gao Q., Chess A.,
 RT Identification of candidate Drosophila olfactory receptors from
 RT genomic DNA sequence.;
 RL Genomics 60:31-39(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Housh K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster";
 Science 287:2185-2195(2000).
 [3]
 RN REVISIONS.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RN IDENTIFICATION, AND TISSUE SPECIFICITY.
 RP MEDLINE=99168688; PubMed=10069338;
 RX Clyne P.J., Warr C.G., Freeman M.R., Lessing D., Kim J., Carlson J.R.;
 RA "A novel family of divergent seven-transmembrane proteins: candidate
 RT odorant receptors in Drosophila";
 RL Neuron 22:327-338(1999).
 [5]
 RN CONCEPTUAL TRANSLATION.
 RP Robertson H.M.;
 RA Unpublished observations (MAY-2003).
 RL -!- FUNCTION: Probable role in the odorant response, being an odorant
 CC receptor.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=a;
 CC IsoId=P81919-1; Sequence=Displayed; CC
 CC -!- Note:No experimental confirmation available;
 CC Name=b;
 CC IsoId=Q9V3N2-1; Sequence=External;
 CC Note:No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Isoform a is expressed in a subset of 17
 CC olfactory receptor neurons in the maxillary palp.
 CC -!- SIMILARITY: Belongs to family Dr-or of G-protein coupled
 CC receptors.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE003830; AAF58834.2; ALT_SEQ.
 CC FlyBase; F59N002388; Or46a.
 CC GO; GO:0016021; C:integral to membrane; NAS.
 CC GO; GO:0004984; P:olfactory receptor activity; NAS.
 CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; NAS.
 CC GO; GO:0007608; P:olfaction; NAS.
 CC InterPro; IPR004117; 7tm_6.
 CC Pfam; PF02949; 7tm_6; 1.
 CC Transmembrane; G-protein coupled receptor; Olfaction;
 CC Alternative splicing; Multigene family.
 KW DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 38 58 1 (POTENTIAL).
 FT DOMAIN 59 65 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 66 86 2 (POTENTIAL).
 FT DOMAIN 87 127 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 128 148 3 (POTENTIAL).
 FT DOMAIN 149 170 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 171 191 4 (POTENTIAL).
 FT DOMAIN 192 255 EXTRACELLULAR (POTENTIAL).
 FT

FT TRANSMEM 256 276 5 (POTENTIAL).
 FT DOMAIN 277 287 CYTOSOLASMIC (POTENTIAL).
 FT TRANSMEM 288 308 6 (POTENTIAL).
 FT DOMAIN 309 355 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 356 376 7 (POTENTIAL).
 FT DOMAIN 377 381 CYTOSOLASMIC (POTENTIAL).
 SQ SEQUENCE 381 AA; 44067 MW; 8A2AB426B08F95FA CRC64;
 Query Match 50.0%; Score 44; DB 1; Length 381;
 Best Local Similarity 60.0%; Pred. No. 7.3; Indels 0; Gaps 0;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VLLLRAGFYAVSFSL 15
 ||:|||||:
 Db 264 VLVLTANFYAIVLS 278
 ||:|||||:
 RESULT 3
 S35D HUMAN STANDARD; PRT; 355 AA.
 ID Q9NTN3; Q92548;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter (UDP-GlcA/UDP-GalNAc transporter) (Solute carrier family 35 member D1).
 GN SLC35D1 OR UGTREL7 OR KIAA0260.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A., AND CHARACTERIZATION.
 TISSUE=Uterus;
 MEDLINE=21223049; PubMed=11322953;
 RA Muraoka M., Kawakita M., Ishida N.;
 RX "Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity.";
 RL FEBS Lett. 495:97-93(2001).
 [2]
 SEQUENCE FROM N.A.
 TISSUE=Bone marrow;
 MEDLINE=97191544; PubMed=9039502;
 RX Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y., Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RA "Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain.";
 RL DNA Res. 3:321-329(1996).
 [3]
 SEQUENCE FROM N.A.
 RA Wilson S.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Transports both UDP-glucuronic acid (UDP-GlcA) and UDP-N-acetylgalactosamine (UDP-GalNAc) from the cytoplasm into the endoplasmic reticulum lumen. May participate in glucuronidation and/or chondroitin sulfate biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- SIMILARITY: Belongs to the nucleotide sugar transporter family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AL133320; CAB92090.1; -
 DR EMBL; D87449; BAAL1390.1; ALT_INT.

DR EMBL; AB044343; BAB18586.1; -
 DR Genew; HGNC:20800; SLC35D1.
 DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.
 DR GO; GO:0005463; F:UDP-N-acetylgalactosamine transporter activity; NAS.
 DR GO; GO:0030206; P:chondroitin sulfate biosynthesis; NAS.
 DR GO; GO:0015789; P:UDP-N-acetylgalactosamine transport; NAS.
 KW transport; Sugar transport; Transmembrane; Endoplasmic reticulum.
 FT TRANSMEM 37 59 POTENTIAL.
 FT TRANSMEM 69 88 POTENTIAL.
 FT TRANSMEM 158 177 POTENTIAL.
 FT TRANSMEM 187 205 POTENTIAL.
 FT TRANSMEM 217 239 POTENTIAL.
 FT TRANSMEM 254 276 POTENTIAL.
 FT TRANSMEM 281 303 POTENTIAL.
 FT TRANSMEM 308 330 POTENTIAL.
 SQ SEQUENCE 355 AA; 39240 MW; CAFF15D49605CA27 CRC64;
 Query Match 49.4%; Score 43.5; DB 1; Length 355;
 Best Local Similarity 70.6%; Pred. No. 8.3; Indels 1; Gaps 1;
 Matches 12; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 QY 3 LLRAGFYAV-SFLSVAV 18
 |||||:
 Db 43 LLRAGFYGVSSFLIVV 59
 |||||:
 RESULT 4
 FRIZ DROME STANDARD; PRT; 581 AA.
 ID FRIZ DROME STANDARD; PRT; 581 AA.
 AC P19537; Q9VUE0; Q9VUE2;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Frizzled protein precursor (Frizzled-1) (Gfzl).
 GN FZ OR CGI7697/CG3646.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 SEQUENCE FROM N.A. (ISOFORM I).
 MEDLINE=89159415; PubMed=2493583;
 RX Vinson C.R., Conover S., Adler P.N.;
 RA "A Drosophila tissue polarity locus encodes a protein containing seven potential transmembrane domains.";
 RL Nature 338:263-264(1989).
 [2]
 SEQUENCE FROM N.A. (ISOFORMS I AND II).
 STRAIN=Canton-S; TISSUE=Embryo;
 RX MEDLINE=91060073; PubMed=2174014;
 RA Adler P.N., Vinson C., Park W.J., Conover S., Klein L.;
 RT "Molecular structure of frizzled, a Drosophila tissue polarity gene.";
 RL Genetics 126:401-416(1990).
 [3]
 SEQUENCE FROM N.A.
 STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., B.D., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davis P., Davenport L.B., Davies P., de la Paolis B., Delcher A., Deng Z., Mays A., Mays A., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.:
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RT Science 302:842-846 (2003).

RN [4]
 RN FUNCTION, TISSUE SPECIFICITY, AND INDUCTION.
 RP MEDLINE=20387013; PubMed=10929111;
 RX Takahashi H., Watanabe-Takahashi A., Smith F.W., Blake-Kaliff M.,
 RA Hawkeford M.J., Saito K.;
 RA "The roles of three functional sulphate transporters involved in
 RT uptake and translocation of sulphate in Arabidopsis thaliana.";
 RL Plant J. 23:171-182 (2000).

RN [5]
 RN INDUCTION.
 RP MEDLINE=22402754; PubMed=12514246;
 RX Okama N., Takei K., Sakakibara H., Hayashi H., Yoneyama T.,
 RA Fujiwara T.;
 RA "Regulation of sulfur-responsive gene expression by exogenously
 RT applied cytokinins in Arabidopsis thaliana.";

RL Plant Cell Physiol. 43:1493-1501 (2002).
 CC -1- FUNCTION: Low-affinity H(+)/sulfate cotransporter that may be
 CC involved in the distribution of sulfate from vascular bundles to
 CC the palisade cells of the leaves. Plays a central role in the
 CC regulation of sulfate assimilation.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- TISSUE SPECIFICITY: Expressed in the phloem in roots and in
 CC the phloem of vascular bundles in leaves.
 CC -1- INDUCTION: In leaves by sulfate starvation. Up-regulated after
 CC treatment with zeatin, an exogenous cytokinin.
 CC -1- SIMILARITY: Belongs to the SLC26A/Sulp transporter (TC 2.A.53)
 CC family.

CC -1- SIMILARITY: Contains 1 STAS domain.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous
 CC gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D85416; BAA12811.1; --
 CC EMBL; AB012047; BAA25174.1; --
 CC EMBL; AC009243; AAF17693.1; ALT_SEQ.
 CC EMBL; AY099863; AAM20714.1; --
 CC EMBL; AY074516; AAL67130.2; --
 CC PIR; S74246; S74246.
 CC InterPro; IPR002645; STAS.
 CC InterPro; IPR001902; Sulph_transpt.
 CC Pfam; PF01740; STAS; 1.
 CC Pfam; PF00916; Sulfate_transp; 1.
 CC TIGRfam; TIGR00815; sulp; 1.
 CC PROSITE; PS01130; SLC26A; 1.
 CC PROSITE; PS00801; STAS; 1.
 CC Transport; Symport; Sulfate transport; Transmembrane; Glycoprotein;
 KW Multigene family.

FT DOMAIN 1 91 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 92 112 POTENTIAL.
 FT DOMAIN 113 114 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 115 135 POTENTIAL.
 FT DOMAIN 136 139 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 140 160 POTENTIAL.
 FT DOMAIN 161 171 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 172 192 POTENTIAL.
 FT DOMAIN 193 194 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 195 215 POTENTIAL.
 FT DOMAIN 216 251 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 252 272 POTENTIAL.
 FT DOMAIN 273 277 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 278 298 POTENTIAL.

FT DOMAIN 299 333 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 334 354 POTENTIAL.
 FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 371 391 POTENTIAL.
 FT DOMAIN 392 403 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 404 424 POTENTIAL.
 FT DOMAIN 425 427 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 428 448 POTENTIAL.
 FT DOMAIN 449 463 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 464 484 POTENTIAL.
 FT DOMAIN 485 658 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 659 687 STAS.
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 89 89 K -> R (IN REF. 1; BAA12811).
 FT CONFLICT 115 115 L -> P (IN REF. 3; AAL67130).
 FT CONFLICT 584 584 V -> M (IN REF. 1; BAA25174).
 SQ SEQUENCE 658 AA; 72317 MW; 587A1D0FD5462FCF CRC64;

Query Match 47.7%; Score 42; DB 1; Length 658;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 LLLRAGFVAVSFLSVAVG 19

DB 465 LVLIAAFGVLFASVEIG 482

RESULT 10
 Y359 AQUAE STANDARD; PRT; 356 AA.
 AC 066685;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AQ_359.
 GN AQ_359.

OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OC NCBI_TaxID=63363;
 RN [1]_TaxID=63363;
 RP SEQUENCE FROM N.A.
 RC STRAIN=VFS;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358 (1998).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB000685; AAC06646.1; --
 CC PIR; A70332; A70332.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 7 29 POTENTIAL.
 FT TRANSMEM 49 71 POTENTIAL.
 FT TRANSMEM 91 113 POTENTIAL.
 FT TRANSMEM 270 292 POTENTIAL.
 FT TRANSMEM 299 316 POTENTIAL.
 FT TRANSMEM 329 348 POTENTIAL.
 SQ SEQUENCE 356 AA; 41807 MW; 4E9EA78C38132D05 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 356;
 Best Local Similarity 37.5%; Pred. No. 21;

Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 LIRAGFYAVSFLSVAV 18
ID YQH2 CAEEL STANDARD; PRT; 498 AA.
DB 53 ILPLGFYISFILTSLI 68

RESULT 11

YQH2 CAEEL
ID YQH2 CAEEL STANDARD; PRT; 498 AA.
AC Q09502; (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable G protein-coupled receptor C43C3.2.
GN C43C3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Sulston J.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Not known. Putative receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC MOST SIMILAR TO SOMATOSTATIN RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 247067; CAB87329.1; -
CC PIR; T19901; T19901.
CC WormPep; C43C3.2; CB01524.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm1; 1.
CC PRINTS; PR00237; GPCR_Rhodopsn.
CC PROSITE; PS00237; G-PROTEIN RECF F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN RECF F1_2; 1.
CC Hypothetical protein; G-protein coupled receptor; Transmembrane;
KW Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 73 1 (POTENTIAL).
FT DOMAIN 74 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 2 (POTENTIAL).
FT DOMAIN 115 130 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 131 151 3 (POTENTIAL).
FT DOMAIN 152 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 191 4 (POTENTIAL).
FT DOMAIN 192 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 279 5 (POTENTIAL).
FT DOMAIN 280 313 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 314 334 6 (POTENTIAL).
FT DOMAIN 335 356 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 357 377 7 (POTENTIAL).
FT DOMAIN 378 498 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 446 450 POLY-ASN.
FT CARBOHYD 48 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 498 AA; 57463 MW; 7F13584348FB546 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 498;

Best Local Similarity 61.5%; Pred. No. 29; Mismatches 0; Gaps 0;

Matches 8; Conservative 3; Indels 2; Indels 0; Gaps 0;

QY 2 LIRAGFYAVSFL 14
ID YQH2 CAEEL STANDARD; PRT; 154 AA.
DB 55 VIARTGFLYIDIAIAG 73

RESULT 13

YI23 ARCFU
ID YI23 ARCFU STANDARD; PRT; 154 AA.
AC O28452;
DT 28-FEB-2003 (Rel. 41, Created)

Query Match

Best Local Similarity 45.5%; Score 40; DB 1; Length 123;

Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAV 19
ID YI23 ARCFU STANDARD; PRT; 154 AA.
DB 55 VIARTGFLYIDIAIAG 73

Db 51 LMLIAGFYALLFM 63

RESULT 12

PHF2_RHIME
ID PHF2_RHIME STANDARD; PRT; 123 AA.
AC Q52965;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phaf2 protein.
GN PHAF2 OR R00397 OR SMC00056.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RCR2011 / SU47;
RC MEDLINE=95276304; PubMed=7756693;
RA Kellier M., Roxlau A., Weng W.M., Schmidt M., Quandt J., Niehaus K.,
RA Jording D., Arnold W., Puchler A.;
RT "Molecular analysis of the Rhizobium meliloti mucR gene regulating the
RT biosynthesis of the exopolysaccharides succinoglycan and
RT galactoglucan";
RL Mol. Plant Microbe Interact. 8:267-277 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=1021;
RC MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Goffard T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puchler A., Purnelle B., Rampsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE PHAF/MRPF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L37353; AAA74237.1; -
CC EMBL; AL591785; CAC45569.1; -
CC InterPro; IPR007208; Mrpf_Phaf.
CC Pfam; PF04066; Mrpf_Phaf; 1.
CC Transport; Potassium transport; Transmembrane; Complete proteome.
FT TRANSMEM 4 24 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
SQ SEQUENCE 123 AA; 12968 MW; 86672D8DB43D0A6E CRC64;

Query Match 45.5%; Score 40; DB 1; Length 123;
Best Local Similarity 36.8%; Pred. No. 12;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAV 19
ID YI23 ARCFU STANDARD; PRT; 154 AA.
DB 55 VIARTGFLYIDIAIAG 73

DR Pfam; PF01275; Myelin_PLP; 1.
DR PRINTS; PRO0214; MYELINPLP.
DR SMART; SM00002; PLP; 1.
DR PROSITE; PS00575; MYELIN_PLP_1; 1.
DR PROSITE; PS01004; MYELIN_PLP_2; 1.
KW Myelin; Transmembrane.
FT TRANSMEM 19 35 POTENTIAL.
FT TRANSMEM 71 87 POTENTIAL.
FT TRANSMEM 117 133 POTENTIAL.
FT TRANSMEM 204 220 POTENTIAL.
SQ SEQUENCE 245 AA; 26910 MW; 521B0A11076B16C5 CRC64;

Query Match 45.5%; Score 40; DB 1; Length 245;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Cy 1 VLLIRAGFYAVSFISVAVG 19
||| | | : |
Db 84 VLLIAGFYTTSAVKALFG 102

Search completed: February 26, 2004, 12:10:49
Job time : 3.61273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 8.11406 Seconds
(without alignments)
738.822 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rv.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	151	6 Q8MJ01	Q8MJ01 macaca mula
2	88	100.0	152	4 Q95668	Q95668 homo sapien
3	88	100.0	165	4 Q95669	Q95669 homo sapien
4	88	100.0	176	4 Q95667	Q95667 homo sapien
5	88	100.0	176	6 Q95J88	Q95J88 macaca fasc
6	88	100.0	177	4 Q14930	Q14930 homo sapien
7	88	100.0	190	4 Q14932	Q14932 homo sapien
8	88	100.0	201	4 Q14931	Q14931 homo sapien
9	85	96.6	180	6 Q8MJ02	Q8MJ02 macaca mula
10	70	79.5	192	11 Q8CG11	Q8CG11 rattus norv
11	70	79.5	192	11 Q8OWM8	Q8OWM8 rattus norv
12	70	79.5	192	11 Q8CFD9	Q8CFD9 rattus norv
13	51	58.0	847	10 Q9SCW1	Q9SCW1 arabidopsis
14	51	58.0	847	10 Q8RWC1	Q8RWC1 arabidopsis
15	49	55.7	610	16 Q7VEF7	Q7VEF7 prochloroc
16	48	54.5	707	10 Q65761	Q65761 cicer ariet

17	48	54.5	721	10 Q9ZP30	Q9ZP30 carica papa
18	47	53.4	838	10 Q9ZP11	Q9ZP11 lycopersico
19	46	52.3	724	10 Q9TOP6	Q9TOP6 lycopersico
20	46	52.3	724	10 Q81100	Q81100 lycopersico
21	46	52.3	757	10 Q80899	Q80899 arabidopsis
22	46	52.3	3763	5 Q8T2A1	Q8T2A1 dictyosteli
23	45	51.1	230	10 Q7X9C6	Q7X9C6 pyrus pyrif
24	45	51.1	270	16 Q8RD77	Q8RD77 thermomane
25	45	51.1	376	10 Q9FVH5	Q9FVH5 prunus arme
26	45	51.1	843	10 Q93X58	Q93X58 fragaria an
27	45	51.1	1561	16 Q888F2	Q888F2 pseudomonas
28	44.5	50.6	479	2 Q86001	Q86001 sphingomonas
29	44	50.0	184	17 Q980V8	Q980V8 sulfolobus
30	44	50.0	351	16 Q8G3L8	Q8G3L8 bifidobacte
31	44	50.0	476	16 Q82WR1	Q82WR1 nitrosomonas
32	44	50.0	591	16 Q8P300	Q8P300 streptococc
33	44	50.0	591	16 Q8K8X7	Q8K8X7 streptococc
34	44	50.0	602	16 Q8CR29	Q8CR29 staphylococ
35	44	50.0	739	10 Q9M5J3	Q9M5J3 phaseolus a
36	44	50.0	845	10 Q9LLS9	Q9LLS9 lycopersico
37	43.5	49.4	202	11 Q8BWB7	Q8BWB7 mus musculu
38	43.5	49.4	306	11 Q8BX24	Q8BX24 mus musculu
39	43.5	49.4	334	11 Q8BXX2	Q8BXX2 mus musculu
40	43	48.9	160	8 Q8M0F0	Q8M0F0 cryptococcu
41	43	48.9	202	8 Q8M0E9	Q8M0E9 cryptococcu
42	43	48.9	202	8 Q85T00	Q85T00 cryptococcu
43	43	48.9	294	17 Q8TQJ7	Q8TQJ7 methanosarc
44	43	48.9	338	17 Q96Y76	Q96Y76 sulfolobus
45	43	48.9	466	16 Q926T9	Q926T9 listeria in

ALIGNMENTS

RESULT 1

Q8MJ01 ID Q8MJ01 PRELIMINARY; PRT; 151 AA.

AC Q8MJ01; PRELIMINARY; PRT; 151 AA.
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE NKP30VIED.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

OC Cercopitheciinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RA LaBonte M.L., Miller J., Letvin N.L.;

RT "Molecular cloning of the rhesus monkey NKP46 and NKP30 and identification

of NKP46SD and NKP30S."

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY035216; AAK63118.1; --

DR InterPro; IPR003599; IG-1.

DR InterPro; IPR007110; IG-1.

DR Pfam; PF00047; IG-1.

DR SMART; SM00409; IG-1.

DR PROSITE; PS00835; IG-1.

SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;

Query Match 100.0%; Score 88; DB 6; Length 151;

Best Local Similarity 100.0%; Pred. No. 1.8e-06; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAVG 19

DB 114 VLLLRAGFYAVSFLSVAVG 132

RESULT 2

Q95668

ID Q95668 PRELIMINARY; PRT; 152 AA.

AC O95669;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7d.
 GN DE
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFkappaB
 RT family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., Demarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670383;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
 RT Level of TNF-beta Production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";

RL Hum. Mol. Genet. 3:793-799(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vainan M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
 RT Nuclear RNA Helicase of the D-E-A-D Family.";
 RL Genomics 26:210-218(1995).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT LKBL and MICA genes at the centromeric end of the HLA Class I
 RT region.";
 RL Genomics 47:372-382(1998).
 DR EMBL; Y14768; CAA75067.1; .
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 SQ SEQUENCE 152 AA; 16393 MW; 42718746451P9ADC CRC64;
 Query Match 100.0%; Score 88; DB 4; Length 152;
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLLLRAGFVAVSFLSVAVG 19
 DB 114 VLLLRAGFVAVSFLSVAVG 132
 RESULT 3
 O95669 PRELIMINARY; PRT; 165 AA.
 ID O95669
 AC O95669;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7f.
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,

RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the NFKappaB
RL family within a 90 kilobase HLA class III segment.";
RN Nat. Genet. 3:137-145(1993).
RP [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8623902;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
RL molecule expressed in transplanted human hearts.";
RN Transplantation 61:1387-1392(1996).
RP [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RL TNF region.";
RN Immunogenetics 42:315-322(1995).
RP [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208881; PubMed=7916555;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
RT "Lymphotoxin-beta: A new member of the TNF family that forms a
RL heteromeric complex with lymphotoxin on the cell surface.";
RN Cell 72:847-856(1993).
RP [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=86016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pemica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotoxin and tumor necrosis factor genes: structure,
RL homology and chromosomal localization.";
RN Nucleic Acids Res. 13:6361-6373(1985).
RP [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=91086846; PubMed=1670638;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Riethmuller G., Weiss E.H.;
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
RL Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
RT Level of TNF-beta Production.";
RL J. Exp. Med. 173:209-219(1991).
RP [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671667;
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RT "Haplotypic polymorphisms of the TNFB gene.";
RL Immunogenetics 33:50-53(1991).
RP [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94362679; PubMed=8081366;
RA Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
RT histocompatibility complex that encodes a potential new member of the
RL I kappa B family of proteins.";
RN Hum. Mol. Genet. 3:793-799(1994).
RP [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324911; PubMed=7601445;
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
RL Nuclear RNA Helicase of the D-E-A-D Family.";
RN Genomics 26:210-218(1995).
RP [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;

RT "Alternative splicing of the LST-1 gene located in the major
RL histocompatibility complex on human chromosome 6.";
RN DNA Seq. 8:155-160(1997).
RP [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=96035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RL differential regulation, initiation, and alternative splicing.";
RN Genomics 45:591-600(1997).
RP [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;
RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
RL Ikbl and MICA genes at the centromeric end of the HLA Class I
RL region.";
RN Genomics 47:372-382(1998).
RP [13]
RP EMBL; Y14768; CAA75068.1; -
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; IG-like.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS00835; IG LIKE; 1.
SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;
Query Match 100.0%; Score 88; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFLSVAVG 19
DB 114 VLLLRAGFYAVSFLSVAVG 132
RESULT 4
O95667 PRELIMINARY; PRT; 176 AA.
AC O95667;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7e.
GN 1C7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the NFKappaB
RL family within a 90 kilobase HLA class III segment.";
RN Nat. Genet. 3:137-145(1993).
RP [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8623902;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
RL molecule expressed in transplanted human hearts.";
RN Transplantation 61:1387-1392(1996).
RP [3]
RP SEQUENCE FROM N.A.

RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
TNF region";
RL Immunogenetics 42:315-322(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208881; PubMed=7916555;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
RT "Lymphotoxin-beta: A new member of the TNF family that forms a
heteromeric complex with lymphotoxin on the cell surface";
RL Cell 72:847-856(1993).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=96016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.D., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotoxin and tumor necrosis factor genes: structure,
homology and chromosomal localization";
RL Nucleic Acids Res. 13:6361-6373(1985).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=91086846; PubMed=1670638;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Riethmüller G., Weiss E.H.;
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
Correlates with A Variant Amino Acid in Position 26 and a Reduced
Level of TNF-beta Production";
RL J. Exp. Med. 173:209-219(1991).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671667;
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RT "Haplotypic polymorphisms of the TNFB gene";
RL Immunogenetics 33:50-53(1991).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=94362679; PubMed=9081366;
RA Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
histocompatibility complex that encodes a potential new member of the
I kappa B family of proteins";
RL Hum. Mol. Genet. 3:793-799(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324911; PubMed=7601445;
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
Nuclear RNA Helicase of the D-E-A-D Family";
RL Genomics 26:210-218(1995).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RT "Alternative splicing of the LST-1 gene located in the major
histocompatibility complex on human chromosome 6";
RL DNA Seq. 8:153-160(1997).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Mater S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
differential regulation, initiation, and alternative splicing";
RL Genomics 45:591-600(1997).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;

RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
IKBL and MICA genes at the centromeric end of the HLA Class I
region";
RL Genomics 47:372-382(1998).
DR EMBL: Y14768; CAA75066.1; --
DR GO: GO:0003793; P:defense/immunity protein activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR GO: GO:0006954; P:inflammatory response; NAS.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
SQ SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;
Query Match 100.0%; Score 88; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFSLVAVG 19
Db 114 VLLLRAGFYAVSFSLVAVG 132
Q95JB8 PRELIMINARY; PRT; 176 AA.
AC Q95JB8; 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE NKP30 (NKP30v1).
GN NCR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541, 9544;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.fascicularis; TISSUE=Lymphoid;
RA Rizzi M., Biasoni R.;
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in
Macaca fascicularis lymphoid cells";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.mulatta;
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification
of NKP46SD and NKP30S";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ278389; CAC41081.1; --
DR EMBL: AY035215; RAK63117.1; --
DR InterPro: IPR007110; Ig-like.
DR Pfam: PF00047; Ig; 1.
DR PROSITE: PS50835; IG LIKE; 1.
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;
Query Match 100.0%; Score 88; DB 6; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFSLVAVG 19
Db 139 VLLLRAGFYAVSFSLVAVG 157

RESULT 6
 ID O14930 PRELIMINARY; PRT; 177 AA.
 AC O14930;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE IC7 precursor (IC7 protein).
 GN IC7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96422187; PubMed=894804;
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";
 RL Genomics 31:215-222(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NF kappa B
 family within a 90 Kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierrina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a
 heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.D., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670639;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced

level of TNF-beta production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
 RA Van Zeveren A., Coppiegers M., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family.";
 RL Genomics 26:210-218(1995).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human Chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IkBL and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 RN [15]
 RP SEQUENCE FROM N.A.
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RP SEQUENCE FROM N.A.
 RA Shiina S., Tamiya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of 5p21.3 HLA class I region.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031136; AAB86578.1; -.
 DR EMBL; Y14768; CAA75064.1; -.
 DR EMBL; AF000505; BAB63393.1; -.
 DR Genew; HGNC:14189; LST1.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0003793; F:defense/immunity protein activity; TAS.
 DR GO; GO:0006968; P:cellular defense response; TAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Signal.

FT SIGNAL 1 61 POTENTIAL.
 FT CHAIN 62 177 1C7.
 SQ SEQUENCE 177 AA; 19237 MW; DD5EC96F0AB2DCE6 CRC64;
 Query Match 100.0%; Score 88; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLLLRAGFYAVSFSLVAVG 19
 |||||
 Db 139 VLLLRAGFYAVSFSLVAVG 157

RESULT 7
 OL14932 PRELIMINARY; PRT; 190 AA.
 ID OL14932 AC OL14932;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7 precursor (Natural killer cell receptor).
 GN 1C7 OR NR230.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96422187; PubMed=8824804;
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."
 RL Genomics 31:215-222(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bouqueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NF kappa B
 RT family within a 90 Kilobase HLA class III segment."
 RL Nat. Genet. 3:137-145(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcoci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts."
 RL Transplantation 61:1387-1392(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region."
 RL Immunogenetics 42:315-322(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface."
 RL Cell 72:847-856(1993).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;

RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization."
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced
 RT level of TNF-beta production."
 RL J. Exp. Med. 173:209-219(1991).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene."
 RL Immunogenetics 33:50-53(1991).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins."
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nures M., Renard C., Geffroin C., Vainan M.,
 RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family."
 RL Genomics 26:210-218(1995).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6."
 RL DNA Seq. 8:155-160(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing."
 RL Genomics 45:591-600(1997).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT LkBL and MICA genes at the centromeric end of the HLA class I
 RT region."
 RL Genomics 47:372-382(1998).
 RN [15]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphoid;
 RA Biassoni R., Pessino A., Malaspina A.;
 RT "NK-A1 activating NK receptor."
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peripheral blood;

RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing";
RL Genomics 45:591-600(1997).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;
RA Shih T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
RT IKB and MICA genes at the centromeric end of the HLA class I
RT region";
RL Genomics 47:372-382(1998).
RN [16]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Wuzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [17]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031137; AAB86579.1; -;
DR EMBL; AF129756; AAD18088.1; -;
DR EMBL; Y14768; CAA75063.1; -;
DR EMBL; BC052582; AAH52582.1; -;
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;
Query Match 100.0%; Score 88; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.4e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFLSVAVG 19
DB 139 VLLLRAGFYAVSFLSVAVG 157
|||||

QSMJ02
ID QSMJ02 PRELIMINARY; PRT; 180 AA.
AC QSMJ02;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE NKp30.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of the rhesus monkey NKp46 and NKp30 and identification
RT of NKp46SD and NKp30S";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035214; AAK63116.1; -;
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;
Query Match 96.6%; Score 85; DB 6; Length 180;
Best Local Similarity 94.7%; Pred. No. 6.7e-06; Mismatches 1; Indels 0; Gaps 0;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFLSVAVG 19
DB 139 VLLLRAGFYAVSFLSVAVG 157
|||||

RESULT 10
Q8CG11 PRELIMINARY; PRT; 192 AA.
ID Q8CG11;
AC Q8CG11; 2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE NK receptor 1c7 precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN;
RA Backman-Petersen E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor 1c7";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430419; CAD23067.2; -;
DR EMBL; AJ430420; CAD23067.2; JOINED.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal; Receptor.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3AFBE6DC0 CRC64;
Query Match 79.5%; Score 70; DB 11; Length 192;
Best Local Similarity 83.3%; Pred. No. 0.002; Mismatches 2; Indels 0; Gaps 0;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 LLLLRAGFYAVSFLSVAVG 19
|||||

```

Db      148 LLLRAGVYALSFSLSVATG 165

RESULT 11
Q8QWMS      PRELIMINARY;      PRT;      192 AA.
AC      Q8QWMS;
DT      01-JUN-2003 (T-EMBLrel. 24, Created)
DT      01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      NRP30.
GN      Rattus norvegicus (Rat).
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LEW;
RX      MEDLINE=22168111; PubMed=12180816;
RA      Hsieh C.L., Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RT      "NK cells and transplantation.";
RT      NK cells and transplantation.";
RL      Transpl. Immunol. 9:111-114(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LEW;
RA      Haieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RT      Martinez O.M., Krams S.M.;
RT      "Identification, Cloning, and Characterization of a Novel Rat NK
RT      Receptor, rNKP30: a Molecule Expressed in Liver Allografts.";
RL      Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY273824; AP13457.1; -.
DR      InterPro; IPR003599; IG.
DR      InterPro; IPR007110; IG-like.
DR      SMART; SM00409; IG; 1.
DR      PROSITE; PS50835; IG LIKE; 1.
SQ      SEQUENCE      192 AA; 20498 MW; 7DF58B245C52377 CRC64;

Query Match      79.5%; Score 70; DB 11; Length 192;
Best Local Similarity      83.3%; Pred. No. 0.002;
Matches      15; Conservative      1; Mismatches      2; Indels      0; Gaps      0;

QY      2 LLLRAGFYAVSFSLSVAVG 19
      ||||| ||||| ||||| |||||
Db      148 LLLRAGVYALSFSLSVATG 165

RESULT 12
Q8CFD9      PRELIMINARY;      PRT;      192 AA.
AC      Q8CFD9;
DT      01-MAR-2003 (T-EMBLrel. 23, Created)
DT      01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      1c7 protein precursor.
GN      1c7.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PVG; TISSUE=NK cell;
RC      Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA      Butcher G.W.;
RT      "Molecular characterization of the novel rat NK receptor 1c7.";
RL      Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ430418; CAD23066.1; -.
DR      InterPro; IPR003599; IG.
DR      InterPro; IPR007110; IG-like.
DR      Pfam; PF00047; IG; 1.
DR      SMART; SM00409; IG; 1.
DR      PROSITE; PS50835; IG LIKE; 1.

RESULT 13
Q9SCW1      PRELIMINARY;      PRT;      847 AA.
AC      Q9SCW1;
DT      01-MAY-2000 (T-EMBLrel. 13, Created)
DT      01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT      01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE      Putative beta-galactosidase precursor (EC 3.2.1.23)
DE      (Lactase).
GN      BGAL1
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Gy I., Kreis M., Lecharny A.;
RT      "the Beta-galactosidases are encoding by a multigene family in
RT      Arabidopsis thaliana.";
RL      Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=COLUMBIA;
RA      Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=COLUMBIA;
RX      PubMed=10907853;
RA      Nakamura Y.;
RT      "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT      Sequence features of the regions of 4,251,695 bp covered by ninety Pl,
RT      TAC and BAC clones.";
RL      DNA Res. 7:217-221(2000).
CC      -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC      GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
CC      -1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
DR      EMBL; AJ270297; CAB64737.1; -.
DR      EMBL; AP001307; BAB01923.1; -.
DR      GO; GO:0009341; C:beta-galactosidase complex; IEA.
DR      GO; GO:0004565; F:beta-galactosidase activity; IEA.
DR      GO; GO:0016738; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR      GO; GO:0004601; F:peroxidase activity; IEA.
DR      GO; GO:0005529; F:sugar binding; IEA.
DR      GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR      GO; GO:0006979; P:response to oxidative stress; IEA.
DR      InterPro; IPR008979; Gal_bind_like.
DR      InterPro; IPR000922; Gal_lectin.
DR      InterPro; IPR001944; Glyco_hydro_35.
DR      InterPro; IPR02016; Peroxidase.
DR      Pfam; PF02140; Gal_Lectin; 1.
DR      Pfam; PF01301; Glyco_hydro_35; 1.
DR      PRINTS; PR00742; GLHYDLASE35.
DR      ProDom; PD005612; Gal_lectin; 1.
DR      PROSITE; PS01182; GLYCOSYL HYDROL F35; 1.
DR      PROSITE; PS00435; PEROXIDASE 1; 1.
DR      PROSITE; PS50228; SUEL LECTIN; 1.
DR      PROSITE; PS50835; Hydrolase; Signal.
KW      Glycosidase, Hydrolase, Signal.

```

```

FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 847 PUTATIVE BETA-GALACTOSIDASE.
SQ SEQUENCE 847 AA; 93658 MW; 91C13DE26A4CF4AD CRC64;

Query Match
Best Local Similarity 58.0%; Score 51; DB 10; Length 847;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19
DB 539 LRAGFNKIALISVAVG 554

RESULT 14
Q8RWCL PRELIMINARY; PRT; 847 AA.
AC Q8RWCL/2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Galactosidase, putative (EC 3.2.1.23) (Beta-galactosidase)
DE (lactase).
GN AT3G33750.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banu J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC -1- GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
DR EMBL; AY093197; AM13196.1; -
DR GO; GO:0009341; C:beta-galactosidase complex; IEA.
DR GO; GO:0004565; F:beta-galactosidase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR008979; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000922; Gal lectin.
DR InterPro; IPR001944; Glyco_hydro_35.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF02140; Gal_Lectin; 1.
DR Pfam; PF01301; Glyco_Hydro_35; 1.
DR PRINTS; PR00742; GLHYDLASE35.
DR PRODOM; PD005612; Gal lectin; 1.
DR PROSITE; PS01182; GLYCOSYL HYDROL_F35; 1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS02028; SUEL_LLECTIN; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 847 AA; 93672 MW; 0F9E12685426CSDA CRC64;

Query Match
Best Local Similarity 58.0%; Score 51; DB 10; Length 847;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19
DB 539 LRAGFNKIALISVAVG 554

RESULT 15
Q7VEF7 PRELIMINARY; PRT; 610 AA.
AC Q7VEF7;
DT 01-OCT-2003 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE ABC-type multidrug transport system ATPase and permease
DE components.
GN MDLB OR PRO00056.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
CX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SARG / CCM 1375 / SS120;
RA MEDLINE=22810154; PubMed=12917486;
RA Dufrene A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RA "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RA a nearly minimal oxyphototrophic genome."
RA Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL; AE017161; AAP99102.1; -
DR Complete proteome.
KW SEQUENCE 610 AA; 68701 MW; EBF39690C516DCDA CRC64;

Query Match
Best Local Similarity 55.7%; Score 49; DB 16; Length 610;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAV 18
DB 156 ILQIVSGFFIVSFISIAI 173

Search completed: February 26, 2004, 12:13:45
Job time : 11.1141 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 11.7931 Seconds
(without alignments)
455.215 Million cell updates/sec

Title: US-10-036-444-5
Perfect score: 88
Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	19	4	AAE02772 Human Nkp
2	88	100.0	177	2	AAV06402 Human B-C
3	88	100.0	190	2	AAV06401 Human B-C
4	88	100.0	190	4	AAE02769 Human Nkp
5	88	100.0	201	4	AAV06403 Human B-C
6	80	90.9	19	5	AAE23900 Human Nkp
7	51	58.0	274	3	AAE23590 Arabidops
8	51	58.0	279	3	AAE23589 Arabidops
9	48	54.5	665	6	AAO16046 Carica pa
10	48	54.5	665	6	AAE32202 Papaya be
11	48	54.5	721	6	AAO16048 Carica pa
12	48	54.5	721	6	AAE32204 Papaya be
13	47	53.4	835	2	AAV44303 Tomato be
14	47	53.4	838	2	AAE2882 Tomato ex
15	47	53.4	838	3	AAV44305 Tomato be
16	46	52.3	114	3	AAE05624 Arabidops
17	46	52.3	119	3	AAE05623 Arabidops
18	46	52.3	127	3	AAE05622 Arabidops
19	46	52.3	661	3	AAE4806 Arabidops
20	46	52.3	714	3	AAE2605 Arabidops
21	46	52.3	724	3	AAV44306 Tomato be
22	46	52.3	757	5	AAE42604 Arabidops
23	46	52.3	757	5	ABE91980 Herbicida
24	45	51.1	1563	6	ABU41824 Protein e
25	44	50.0	87	7	ADC97632 E. faeciu

RESULT 1
AAE02772
ID AAE02772 standard; peptide; 19 AA.
XX
AC AAE02772;
XX
DT 06-AUG-2001 (first entry)
XX
DE Human Nkp30 receptor transmembrane region sequence.
XX
KW Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;
KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
KW therapy; transmembrane region.
XX
OS Homo sapiens.
XX
PN WO200136630-A2.
XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-EP011697.
XX
PR 15-NOV-1999; 99CA-02288307.
XX
PR 15-NOV-1999; 99US-00440514.
XX
PA (INNA-) INNATE PHARMA SAS.
XX
PA (UYGE-) UNIV GENOVA.
XX
PI Moretta A, Bottino C, Biassoni R;
XX
DR MPI; 2001-329221/34.
XX
PT Novel compound, useful for detection and/or quantifying the presence of
XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.
XX
PF Claim 1; Fig 7B; 83pp; English.

ALIGNMENTS

26	44	50.0	385	3	AAE20910
27	44	50.0	385	7	ADD15306 Fruitfly
28	44	50.0	417	4	AG81928 S. epider
29	44	50.0	425	5	ABP6312 Bifidobac
30	44	50.0	602	5	ABP39571 Staphyloc
31	44	50.0	602	6	ABU42891 Protein e
32	43.5	49.4	383	6	ABO00770 Polypepti
33	43	48.9	154	6	ABO00816 Polypepti
34	43	48.9	166	4	ABE66413 Drosophil
35	43	48.9	469	4	ABE76682 Corynebac
36	43	48.9	530	4	ABE76681 Corynebac
37	43	48.9	530	4	ABE79761 Corynebac
38	43	48.9	578	4	AAE90808 C Glutami
39	43	48.9	589	4	ABE71245 Drosophil
40	42	47.7	271	6	ABU35719 Protein e
41	42	47.7	466	5	ABE47399 Listeria
42	41.5	47.2	588	6	AAE32079 Human TRI
43	41.5	47.2	589	5	AAU9329 Human tra
44	41.5	47.2	589	6	AAO30994 Human tra
45	41.5	47.2	589	7	ADD01392 Human TCH

The invention relates to human Nkp30 receptor and its corresponding cDNA molecule which is involved in natural cytotoxicity mediated by natural killer (NK) cells and antibodies that identify the same. Nkp30 is selectively is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively expressed on the surface of human mature NK cells. Nkp30 and its cDNA are useful for detecting and/or quantifying the presence of NK cells in a biological sample. The invention also provide kits for detecting and/or quantifying the presence of NK cells, for the selective removal of NK cells from a biological sample, for the positive and selective purification of NK cells from a biological sample and for the in vitro

stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GvT) and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. Nkp30 antibodies are useful for identifying Nkp30 natural ligands and allow assessment of the level of surface Nkp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence Nkp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the transmembrane region of human Nkp30 receptor

Query Match 100.0%; Score 88; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.8e-08; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 VLLLRAGFYAVSFSLVAVG 19
Db 1 VLLLRAGFYAVSFSLVAVG 19

RESULT 2
AAV06402
ID AAY06402 standard; protein; 177 AA.

XX AC AAY06402;
XX DT 20-SEP-1999 (first entry)

XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
XX KW signal transduction; immunomodulator; antiinflammatory;
XX KW autoimmune disease; inflammation; gene therapy; diagnosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Peptide 1..12 /note= "leader peptide"
XX FT Protein 13..177 /note= "mature protein"
XX FT Modified-site 42 /note= "N-glycosylated"
XX FT Modified-site 68 /note= "N-glycosylated"
XX FT Modified-site 121 /note= "N-glycosylated"
XX FT Domain 139..162 /note= "transmembrane domain"
XX FT Peptide 166..177 /note= "alternatively spliced C-terminal end"

XX PN WO9923867-A2.
XX XX 20-MAY-1999.
XX PF 05-NOV-1998; 98WO-US023826.
XX PR 07-NOV-1997; 97US-0064761P.
XX PA (BIOJ) BIOGEN INC.
XX PI Browning J;
XX XX WPI; 1999-418423/35.
XX DR N-PSDB; AAX59348.

XX PT Novel B-cell myelin oligodendrocyte glycoproteins.
XX DR

PS Claim 2; Page 43; 43pp; English.
XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

XX SQ Sequence 177 AA;

Query Match 100.0%; Score 88; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 VLLLRAGFYAVSFSLVAVG 19
Db 139 VLLLRAGFYAVSFSLVAVG 157

RESULT 3
AAV06401
ID AAY06401 standard; protein; 190 AA.

XX AC AAY06401;
XX XX 20-SEP-1999 (first entry)

XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
XX KW signal transduction; immunomodulator; antiinflammatory;
XX KW autoimmune disease; inflammation; gene therapy; diagnosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT Peptide 1..12 /note= "leader peptide"
XX FT Protein 13..190 /note= "mature protein"
XX FT Modified-site 42 /note= "N-glycosylated"
XX FT Modified-site 68 /note= "N-glycosylated"
XX FT Modified-site 121 /note= "N-glycosylated"
XX FT Domain 139..162 /note= "transmembrane domain"
XX FT Peptide 166..190 /note= "alternatively spliced C-terminal end"

XX PN WO9923867-A2.
XX XX 20-MAY-1999.
XX PF 05-NOV-1998; 98WO-US023826.
XX PR 07-NOV-1997; 97US-0064761P.
XX PA (BIOJ) BIOGEN INC.
XX PI Browning J;
XX XX WPI; 1999-418423/35.
XX DR

DR N-PSDB; AAX59347.
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 PT Claim 2; Page 42; 43pp; English.
 XX
 XX This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and is soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 SQ Sequence 190 AA;
 Query Match 100.0%; Score 88; DB 2; Length 190;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0;
 QY 1 VLLLRAGFYAVSFSLVAVG 19
 DB 139 VLLLRAGFYAVSFSLVAVG 157
 RESULT 4
 AAE02769
 ID AAE02769 standard; protein; 190 AA.
 XX
 AC AAE02769;
 XX
 DT 06-AUG-2001 (first entry)
 XX
 DE Human NKp30 receptor.
 XX
 KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..18 /label= Signal_peptide
 FT Protein 19..190 /label= Mature_NKp30_receptor_protein
 FT Region 19..138 /label= Extracellular_region
 FT Modified-site 42 /note= "Forms an immunoglobulin (Ig) V-like domain"
 FT Modified-site 121 /note= "N-glycosylation site"
 FT Modified-site 139..157 /note= "N-glycosylation site"
 FT Region 158..190 /label= Transmembrane_region
 FT Region /label= Intracellular_region
 XX
 DN WO200136630-A2.
 XX
 PD 25-MAY-2001.
 XX
 XX 15-NOV-2000; 2000WO-EP011697.
 PF
 XX 15-NOV-1999; 99CA-02288307.
 PR

PR 15-NOV-1999; 99US-00440514.
 XX (INNA-) INNATE PHARMA SAS.
 PA (UYGE-) UNIV GENOVA.
 XX
 PI Moretta A, Bottino C, Biassoni R;
 XX
 DR WPI; 2001-329221/34.
 DR N-PSDB; AAD06564.
 XX
 PT Novel compound, useful for detection and/or quantifying the presence of
 FT NK cells, comprises the amino acid sequences of the NKp30 molecule.
 XX
 XX Claim 1; Fig 7B; 83pp; English.
 XX
 XX The invention relates to human NKp30 receptor and its corresponding cDNA
 CC molecule which is involved in natural cytotoxicity mediated by natural
 CC killer (NK) cells and antibodies that identify the same. NKp30 receptor
 CC is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively
 CC expressed on the surface of human mature NK cells. NKp30 and its cDNA are
 CC useful for detecting and/or quantifying the presence of NK cells in a
 CC biological sample. The invention also provides kits for detecting and/or
 CC quantifying the presence of NK cells, for the selective removal of NK
 CC cells from a biological sample, for the positive and selective
 CC purification of NK cells from a biological sample and for the in vitro
 CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GvH) inhibition, stimulation of graft
 CC versus tumour (GvT) and especially graft versus leukaemia (GvL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. NKp30 antibodies are useful for
 CC identifying NKp30 natural ligands and allow assessment of the level of
 CC surface NKp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence NKp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is human NKp30 receptor
 XX
 SQ Sequence 190 AA;
 Query Match 100.0%; Score 88; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 6.8e-07; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0;
 QY 1 VLLLRAGFYAVSFSLVAVG 19
 DB 139 VLLLRAGFYAVSFSLVAVG 157
 RESULT 5
 AAY06403
 ID AAY06403 standard; protein; 201 AA.
 XX
 AC AAY06403;
 XX
 DT 20-SEP-1999 (first entry)
 XX
 DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 XX
 KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
 KW signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..12 /note= "leader peptide"
 FT Protein 13..201 /note= "mature protein"
 FT Modified-site 42 /note= "N-glycosylated"
 FT Modified-site 68

FT Modified-site /note= "N-glycosylated"
 FT 121
 FT Domain /note= "N-glycosylated"
 FT 139..162
 FT Peptide /note= "transmembrane domain"
 FT 166..201
 FT /note= "alternatively spliced C-terminal end"
 XX
 PN WO9923867-A2.
 XX
 XX 20-MAY-1999.
 XX
 XX 05-NOV-1998; 98WO-US023826.
 XX
 XX 07-NOV-1997; 97US-0064761P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 XX Browning J;
 XX
 XX WPI; 1999-418423/35.
 DR N-PSDB; AAX59349.
 XX
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 FT
 XX
 XX Claim 2; Page 43; 43pp; English.
 PS
 XX This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AA06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 SQ Sequence 201 AA;
 Query Match 100.0%; Score 88; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 7.2e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLLLRAGFYAVSFSLVAVG 19
 DB 139 VLLLRAGFYAVSFSLVAVG 157
 RESULT 6
 AAEE23900
 ID AAEE23900 standard; peptide; 19 AA.
 XX
 AC AAEE23900;
 XX
 DT 10-SEP-2002 (first entry)
 XX
 DE Human Nkp30 receptor peptide.
 XX
 KW KAR-associated protein; KARAP-transduced immune signal; dendritic cell;
 KW antigen presentation; contact sensitivity; multiple sclerosis;
 KW neuroprotective; human; Nkp30 receptor peptide.
 XX
 OS Homo sapiens.
 XX
 XX WO200224940-A2.
 PN
 XX 28-MAR-2002.
 PD
 XX

PF 20-SEP-2001; 2001WO-EP011492.
 XX
 PR 20-SEP-2000; 2000US-0234161P.
 XX
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX
 XX Vivier E, Vely F, Tomasello E;
 PI
 XX WPI; 2002-454420/48.
 DR
 XX Identifying KAR-associated protein-transduced immune signal inhibitor,
 PT comprises using cells co-expressing functional KARAP, and engineered
 PT cells and animals that over-express functional KARAP or bear non-
 PT functional KARAP.
 XX
 XX Example 4; Page 45; 89pp; English.
 PS
 XX The present invention relates to a novel method for identifying compounds
 CC capable of inhibiting KAR-associated protein (KARAP)-transduced immune
 CC signals. The method involves using functional and non-functional KARAP,
 CC cells co-expressing functional KARAP, functional receptors transducing
 CC their signal by zeta, gamma or epsilon, and engineered cells and animals
 CC over-expressing functional KARAP or bearing non-functional KARAP. The
 CC method is useful for identifying compounds capable of inhibiting KARAP-
 CC transduced immune signals. The KARAP-inhibiting compounds are useful for
 CC impairing the development and maturation of dendritic cells, for
 CC inhibiting the antigen presentation of dendritic cells, by synthesis
 CC inhibition or through inhibition of the migration of dendritic cells, for
 CC making drugs intended for inhibiting dendritic cell development or
 CC maturation, for preparing drugs for the treatment, prevention, palliation
 CC of immune response, where the activation of KAR has to be inhibited and
 CC for the treatment of contact sensitivity or multiple sclerosis. The
 CC present sequence is human Nkp30 receptor peptide which associate with CD3
 CC zeta and FcR gamma. This sequence is used in the exemplification of the
 CC invention
 XX
 SQ Sequence 19 AA;
 Query Match 90.9%; Score 80; DB 5; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LLRAGFYAVSFSLVAVG 19
 DB 1 LLRAGFYAVSFSLVAVG 17
 RESULT 7
 AAG23590
 ID AAG23590 standard; protein; 274 AA.
 XX
 AC AAG23590;
 XX
 XX 17-OCT-2000 (first entry)
 DT
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 26954.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 XX BP1033405-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-00301439.
 PF
 XX 25-FEB-1999; 99US-0121825P.
 PR
 XX 05-MAR-1999; 99US-0123180P.
 PR
 XX 09-MAR-1999; 99US-0123548P.
 PR
 XX 23-MAR-1999; 99US-0125788P.
 PR

PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 03-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149732P.
 PR 20-AUG-1999; 99US-0149923P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.

PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159337P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 14-OCT-1999; 99US-0159683P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 21-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 58.0%; Score 51; DB 3; Length 274;

Best Local Similarity 62.5%; Pred. No. 2.1;

Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 4 LRAGFYAVSFSLVAVG 19
 ||||| : : : : :
 Db 30 LRAGFNKIALISIAVG 45

RESULT 8
 AAG23589
 ID AAG23589 standard; protein; 279 AA.

XX AC AAG23589;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 26953.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PP 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-01213180P.

PR 23-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 29-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 08-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137744P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139432P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142053P.
 PR 06-JUL-1999; 99US-0142300P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144332P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149829P.
 PR 23-AUG-1999; 99US-0149802P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159332P.
 PR 14-OCT-1999; 99US-0159333P.
 PR 14-OCT-1999; 99US-0159334P.
 PR 14-OCT-1999; 99US-0159335P.
 PR 14-OCT-1999; 99US-0159336P.
 PR 14-OCT-1999; 99US-0159337P.
 PR 14-OCT-1999; 99US-0159338P.
 PR 14-OCT-1999; 99US-0159339P.
 PR 14-OCT-1999; 99US-0159340P.
 PR 14-OCT-1999; 99US-0159341P.
 PR 14-OCT-1999; 99US-0159342P.
 PR 14-OCT-1999; 99US-0159343P.
 PR 14-OCT-1999; 99US-0159344P.
 PR 14-OCT-1999; 99US-0159345P.
 PR 14-OCT-1999; 99US-0159346P.
 PR 14-OCT-1999; 99US-0159347P.
 PR 14-OCT-1999; 99US-0159348P.
 PR 14-OCT-1999; 99US-0159349P.
 PR 14-OCT-1999; 99US-0159350P.
 PR 14-OCT-1999; 99US-0159351P.
 PR 14-OCT-1999; 99US-0159352P.
 PR 14-OCT-1999; 99US-0159353P.
 PR 14-OCT-1999; 99US-0159354P.
 PR 14-OCT-1999; 99US-0159355P.
 PR 14-OCT-1999; 99US-0159356P.
 PR 14-OCT-1999; 99US-0159357P.
 PR 14-OCT-1999; 99US-0159358P.
 PR 14-OCT-1999; 99US-0159359P.
 PR 14-OCT-1999; 99US-0159360P.
 PR 14-OCT-1999; 99US-0159361P.
 PR 14-OCT-1999; 99US-0159362P.
 PR 14-OCT-1999; 99US-0159363P.
 PR 14-OCT-1999; 99US-0159364P.
 PR 14-OCT-1999; 99US-0159365P.
 PR 14-OCT-1999; 99US-0159366P.
 PR 14-OCT-1999; 99US-0159367P.
 PR 14-OCT-1999; 99US-0159368P.
 PR 14-OCT-1999; 99US-0159369P.
 PR 14-OCT-1999; 99US-0159370P.
 PR 14-OCT-1999; 99US-0159371P.
 PR 14-OCT-1999; 99US-0159372P.
 PR 14-OCT-1999; 99US-0159373P.
 PR 14-OCT-1999; 99US-0159374P.
 PR 14-OCT-1999; 99US-0159375P.
 PR 14-OCT-1999; 99US-0159376P.
 PR 14-OCT-1999; 99US-0159377P.
 PR 14-OCT-1999; 99US-0159378P.
 PR 14-OCT-1999; 99US-0159379P.
 PR 14-OCT-1999; 99US-0159380P.
 PR 14-OCT-1999; 99US-0159381P.
 PR 14-OCT-1999; 99US-0159382P.
 PR 14-OCT-1999; 99US-0159383P.
 PR 14-OCT-1999; 99US-0159384P.
 PR 14-OCT-1999; 99US-0159385P.
 PR 14-OCT-1999; 99US-0159386P.
 PR 14-OCT-1999; 99US-0159387P.
 PR 14-OCT-1999; 99US-0159388P.
 PR 14-OCT-1999; 99US-0159389P.
 PR 14-OCT-1999; 99US-0159390P.
 PR 14-OCT-1999; 99US-0159391P.
 PR 14-OCT-1999; 99US-0159392P.
 PR 14-OCT-1999; 99US-0159393P.
 PR 14-OCT-1999; 99US-0159394P.
 PR 14-OCT-1999; 99

CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 665 AA;

Query Match 54.5%; Score 48; DB 6; Length 665;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 393 LRAGINKISILSIIVG 398

RESULT 10
AAE32202
ID AAE32202 standard; protein; 665 AA.
XX
AC AAE32202;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).45.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.45; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011803.
XX
PR 11-APR-2001; 2001US-0283022P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A, Chiang C;
XX
XX WPI; 2003-075493/07.
XX
DR N-PSDB; AAD49832.
XX
PT New DNA construct, useful for controlling the ripening of papaya fruit
PT and conferring resistance to papaya ringspot virus coat in transgenic
PT plants.
XX
PS Disclosure; Page 75-77; 121pp; English.

CC The invention relates to a DNA construct comprising a first DNA molecule
CC encoding a protein that controls papaya fruit ripening and a second DNA
CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The
CC DNA construct is useful for controlling the ripening of papaya fruit and
CC conferring resistance to PRSV coat in transgenic plants. The present
CC sequence is papaya beta-galactosidase (beta-Gal).45
XX
SQ Sequence 665 AA;

Query Match 54.5%; Score 48; DB 6; Length 665;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 393 LRAGINKISILSIIVG 398

RESULT 12
AAE32204
ID AAE32204 standard; protein; 721 AA.
XX
AC AAE32204;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).41.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011803.
XX
PR 11-APR-2001; 2001US-0283022P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A, Chiang C;
XX
XX WPI; 2003-075493/07.
XX
DR N-PSDB; AAD49832.
XX
PT New DNA construct, useful for controlling the ripening of papaya fruit
PT and conferring resistance to papaya ringspot virus coat in transgenic
PT plants.
XX
PS Disclosure; Page 75-77; 121pp; English.

CC The invention relates to a DNA construct comprising a first DNA molecule
CC encoding a protein that controls papaya fruit ripening and a second DNA
CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The
CC DNA construct is useful for controlling the ripening of papaya fruit and
CC conferring resistance to PRSV coat in transgenic plants. The present
CC sequence is papaya beta-galactosidase (beta-Gal).45
XX
SQ Sequence 665 AA;

Query Match 54.5%; Score 48; DB 6; Length 665;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 393 LRAGINKISILSIIVG 398

RESULT 12
AAE32204
ID AAE32204 standard; protein; 721 AA.
XX
AC AAE32204;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).41.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011804.
XX
PR 11-APR-2001; 2001US-0283008P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A;
XX
XX WPI; 2003-075556/07.
XX
DR N-PSDB; AAL51074.
XX
PT New isolated nucleic acid molecule encoding beta-galactosidase protein,
PT useful for promoting or delaying papaya fruit ripening.
XX
PS Claim 17; Page 15-19; 84pp; English.

CC The invention comprises the amino acid and coding sequence of Carica
CC papaya beta-galactosidase proteins which control papaya fruit ripening.
CC The invention also comprises the amino acid and coding sequence of Carica
CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and
CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 721 AA;

Query Match 54.5%; Score 48; DB 6; Length 721;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 526 LRAGVKNVLSLSIIVG 541

RESULT 11
AAO16048
ID AAO16048 standard; protein; 721 AA.
XX
AC AAO16048;
XX
DT 27-FEB-2003 (first entry)
XX
DE Carica papaya beta-galactosidase (beta-gal-41).
XX
KW Enzyme; papaya; beta-galactosidase; pectinmethylesterase;
KW polygalacturonase; fruit ripening.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200283924-A2.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011804.
XX
PR 11-APR-2001; 2001US-0283008P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A;
XX
XX WPI; 2003-075556/07.
XX
DR N-PSDB; AAL51074.
XX
PT New isolated nucleic acid molecule encoding beta-galactosidase protein,
PT useful for promoting or delaying papaya fruit ripening.
XX
PS Claim 17; Page 15-19; 84pp; English.

CC The invention comprises the amino acid and coding sequence of Carica
CC papaya beta-galactosidase proteins which control papaya fruit ripening.
CC The invention also comprises the amino acid and coding sequence of Carica
CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and
CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 721 AA;

Query Match 54.5%; Score 48; DB 6; Length 721;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 526 LRAGVKNVLSLSIIVG 541

RESULT 12
AAE32204
ID AAE32204 standard; protein; 721 AA.
XX
AC AAE32204;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).41.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011804.
XX
PR 11-APR-2001; 2001US-0283008P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A;
XX
XX WPI; 2003-075556/07.
XX
DR N-PSDB; AAL51074.
XX
PT New isolated nucleic acid molecule encoding beta-galactosidase protein,
PT useful for promoting or delaying papaya fruit ripening.
XX
PS Claim 17; Page 15-19; 84pp; English.

CC The invention comprises the amino acid and coding sequence of Carica
CC papaya beta-galactosidase proteins which control papaya fruit ripening.
CC The invention also comprises the amino acid and coding sequence of Carica
CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and
CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 721 AA;

Query Match 54.5%; Score 48; DB 6; Length 721;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 526 LRAGVKNVLSLSIIVG 541

RESULT 12
AAE32204
ID AAE32204 standard; protein; 721 AA.
XX
AC AAE32204;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).41.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011804.
XX
PR 11-APR-2001; 2001US-0283008P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A;
XX
XX WPI; 2003-075556/07.
XX
DR N-PSDB; AAL51074.
XX
PT New isolated nucleic acid molecule encoding beta-galactosidase protein,
PT useful for promoting or delaying papaya fruit ripening.
XX
PS Claim 17; Page 15-19; 84pp; English.

CC The invention comprises the amino acid and coding sequence of Carica
CC papaya beta-galactosidase proteins which control papaya fruit ripening.
CC The invention also comprises the amino acid and coding sequence of Carica
CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and
CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 721 AA;

Query Match 54.5%; Score 48; DB 6; Length 721;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 526 LRAGVKNVLSLSIIVG 541

RESULT 12
AAE32204
ID AAE32204 standard; protein; 721 AA.
XX
AC AAE32204;
XX
DT 24-MAR-2003 (first entry)
XX
DE Papaya beta-galactosidase (beta-Gal).41.
XX
KW Fruit ripening; papaya ringspot virus coat protein; transgenic plant;
KW transgenic: PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.
XX
OS Carica papaya.

XX Key Location/Qualifiers
FH Misc-difference 143
FT /label= "Unknown"
FT /note= "Encoded by TGN"
XX
XX WO200282889-A1.
XX
PD 24-OCT-2002.
XX
PF 11-APR-2002; 2002WO-US011804.
XX
PR 11-APR-2001; 2001US-0283008P.
XX
PA (CORR) CORNELL RES FOUND INC.
PA (PAIS/) PAIS M S S.

PI Pais MSS, Gonsalves D, Balde A;
XX
XX WPI; 2003-075556/07.
XX
DR N-PSDB; AAL51074.
XX
PT New isolated nucleic acid molecule encoding beta-galactosidase protein,
PT useful for promoting or delaying papaya fruit ripening.
XX
PS Claim 17; Page 15-19; 84pp; English.

CC The invention comprises the amino acid and coding sequence of Carica
CC papaya beta-galactosidase proteins which control papaya fruit ripening.
CC The invention also comprises the amino acid and coding sequence of Carica
CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and
CC protein sequences of the invention are useful for promoting or delaying
CC papaya fruit ripening. The present amino acid sequence represents a
CC Carica papaya fruit ripening-related protein of the invention
XX
SQ Sequence 721 AA;

Query Match 54.5%; Score 48; DB 6; Length 721;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 526 LRAGVKNVLSLSIIVG 541

XX PD 24-OCT-2002.
 XX PF 11-APR-2002; 2002WO-US011803.
 XX PR 11-APR-2001; 2001US-0283022P.
 XX PS (CORR) CORNELL RES FOUND INC.
 PA (PAIS/) PAIS M S S.
 XX PA Pais MSS, Gonsalves D, Balde A, Chiang C;
 PI N-PSDB; AAD49834.
 XX WPI; 2003-075493/07.
 XX N-PSDB; AAD49834.
 XX New DNA construct, useful for controlling the ripening of papaya fruit
 PT and conferring resistance to papaya ringspot virus coat in transgenic
 PT plants.
 XX PS Disclosure; Page 80-83; 121pp; English.
 XX CC The invention relates to a DNA construct comprising a first DNA molecule
 CC encoding a protein that controls papaya fruit ripening and a second DNA
 CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The
 CC DNA construct is useful for controlling the ripening of papaya fruit and
 CC conferring resistance to PRSV coat in transgenic plants. The present
 CC sequence is papaya beta-galactosidase (beta-Gal).41
 XX SQ Sequence 721 AA;
 Query Match 54.5%; Score 48; DB 6; Length 721;
 Best Local Similarity 68.8%; Pred. No. 21;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 4 LRAGFYAVSFSLVAVG 19
 ||||| :|||:
 DB 526 LRAGVNVKSLLSIAVG 541
 RESULT 13
 AAY44303
 ID AAY44303 standard; protein; 835 AA.
 XX AC AAY44303;
 XX DT 29-FEB-2000 (first entry)
 XX DE Tomato beta galactosidase-1.
 XX KW Tomato beta galactosidase-1; TBG; Rutgers tomato plant; pectin;
 KW fruit softening; beta galactosidase II protein; biofilm;
 KW transgenic plant; protoplast isolation.
 XX OS Lycopersicon esculentum.
 XX FH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label= Signal_peptide
 FT 25..835
 FT /label= beta-galactosidase-1
 XX WO9964564-A1.
 XX PD 16-DEC-1999.
 XX PF 08-JUN-1999; 99WO-US012697.
 XX PR 09-JUN-1998; 98US-0088805P.
 XX PA (USDA) US DEPT OF AGRICULTURE.
 PI Gross KC, Smith DL;
 XX

DR WPI; 2000-097532/08.
 DR N-PSDB; AAZ29338.
 XX New beta-galactosidases, used to prepare transgenic plants with altered
 PT fruit ripening.
 XX PS Claim 1; Fig 2; 85pp; English.
 XX CC The present sequence is tomato beta galactosidase-1 (TBG-1) encoded by a
 CC cDNA derived from breaker, turning and pink fruit pericarp from 'Rutgers'
 CC tomato plants. This hydrolyses terminal non-reducing beta-D-galactosyl
 CC residues from beta-D-galactosides leading to loss of tissue integrity and
 CC fruit softening. This is used for modifying cell wall metabolism and
 CC controlling ripening of fruit by altering activity of beta galactosidase
 CC II protein. Pectin with reduced galactosyl content is produced for use in
 CC biofilms or solutions. Transgenic plants with altered fruit ripening are
 CC produced by introducing DNA constructs comprising TBG cDNA. TBG forms a
 CC component of an enzyme mixture used to isolate protoplasts
 XX SQ Sequence 835 AA;
 Query Match 53.4%; Score 47; DB 3; Length 835;
 Best Local Similarity 62.5%; Pred. No. 37;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 4 LRAGFYAVSFSLVAVG 19
 ||||| :|||:
 DB 527 LRAGVNVKSLLSIAVG 542
 RESULT 14
 AAR82882
 ID AAR82882 standard; protein; 838 AA.
 XX AC AAR82882;
 XX DT 10-FEB-1996 (first entry)
 XX DE Tomato exo-(1-4)beta-D-galactanase enzyme.
 XX KW Exo-(1-4)-beta-D-galactanase; enzyme; transgenic plant; crop improvement.
 XX OS Lycopersicon esculentum.
 XX PN WO9523228-A1.
 XX PD 31-AUG-1995.
 XX PF 23-FEB-1995; 95WO-GB000372.
 XX PR 23-FEB-1994; 94GB-00003423.
 XX PA (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 XX Chengappa S, Hellyer SA, De Silva J, Reid JSG;
 XX WPI; 1995-311537/40.
 DR N-PSDB; AAT01015.
 XX Lupin and tomato exo-galactanase DNA and protein - useful for the
 PT alteration of plant characteristic(s), e.g. texture, growth, ripening.
 XX PS Disclosure; Page 38; 67pp; English.
 XX CC This enzyme may be expressed in a transgenic plant in order to alter
 CC characteristics of e.g. growth, texture or ripening of the plant or plant
 CC parts. This protein may also be expressed in a recombinant host for the
 CC production of the enzyme which may be used for the modification,
 CC degradation or liquefaction of plant materials in order to affect
 CC mechanical properties relating to eating texture, particle sizes of, e.g.
 CC fruit or vegetable juices, or extractability of colors, flavors or
 CC vitamins

```
XX SQ Sequence 838 AA;
Query Match 53.4%; Score 47; DB 2; Length 838;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
|||||:|||||
Db 530 LRAGWNKISLSIAVG 545

Search completed: February 26, 2004, 12:10:10
Job time : 15.9931 secs

RESULT 15
AAY44305
ID AAY44305 standard; protein; 838 AA.
XX AC AAY44305;
XX DT 29-FEB-2000 (first entry)
XX DE Tomato beta galactosidase-3.
XX KW Tomato beta galactosidase-3; TBG; Rutgers tomato plant; pectin;
KW fruit softening; beta galactosidase II protein; biofilm;
KW transgenic plant; protoplast isolation.
XX OS Lycopersicon esculentum.
XX FH Key Location/Qualifiers
FT Peptide 1..22
FT Protein /label= signal_peptide
FT 23..838
FT /label= beta-galactosidase-3
XX PN WO9964564-A1.
XX PD 16-DEC-1999.
XX PF 08-JUN-1999; 99WO-US012697.
XX PR 09-JUN-1998; 98US-0088805P.
XX PA (USDA ) US DEPT OF AGRICULTURE.
XX PI Groes KC, Smith DL;
XX DR WPI; 2000-097532/08.
XX DR N-PSDB; AAZ29340.
XX PT New beta-galactosidases, used to prepare transgenic plants with altered
XX fruit ripening.
XX PS Claim 1; Fig 2; 85pp; English.
XX CC The present sequence is tomato beta galactosidase-3 (TBG-3) encoded by a
CC cDNA derived from breaker, turning and pink fruit pericarp from 'Rutgers'
CC tomato plants. This hydrolyses terminal non-reducing beta-D-galactosyl
CC residues from beta-D-galactosides leading to loss of tissue integrity and
CC fruit softening. TBG3 transcript was detected at low levels in root and
CC stem tissues. This is used for modifying cell wall metabolism and
CC controlling ripening of fruit by altering activity of beta galactosidase
CC II protein. Pectin with reduced galactosyl content is produced for use in
CC biofilms or solutions. Transgenic plants with altered fruit ripening are
CC produced by introducing DNA constructs comprising TBG cDNA. TBG forms a
CC component of an enzyme mixture used to isolate protoplasts
XX SQ Sequence 838 AA;

Query Match 53.4%; Score 47; DB 3; Length 838;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 6.60212 Seconds
(without alignments)
607.670 Million cell updates/sec

Title: US-10-036-444-5
Perfect score: 88
Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	19	13	US-10-036-444-5
2	88	100.0	190	13	Sequence 5, Appli
3	48	54.5	665	14	Sequence 2, Appli
4	48	54.5	665	15	Sequence 2, Appli
5	48	54.5	721	14	Sequence 6, Appli
6	48	54.5	721	15	Sequence 6, Appli
7	44.5	50.6	473	15	Sequence 10510, A
8	43	48.9	536	10	Sequence 42, Appl
9	43	48.9	578	9	Sequence 4562, Ap
10	41.5	47.2	589	9	Sequence 2, Appli
11	41.5	47.2	589	14	Sequence 2, Appli
12	41.5	47.2	850	9	Sequence 3, Appli
13	41	46.6	297	9	Sequence 6265, Ap
14	41	46.6	320	14	Sequence 9631, Ap
15	41	46.6	333	14	Sequence 441, App

16	41	46.6	405	10	US-09-932-227-60
17	41	46.6	405	14	US-10-183-708-60
18	41	46.6	421	14	US-10-156-761-12314
19	40	45.5	97	9	US-09-738-626-3787
20	40	45.5	99	9	US-09-864-761-42102
21	40	45.5	220	15	US-10-369-493-12842
22	40	45.5	257	14	US-10-156-761-13244
23	40	45.5	278	14	US-10-029-386-32250
24	40	45.5	351	10	US-09-882-227-434
25	40	45.5	477	14	US-10-245-537A-4
26	40	45.5	551	13	US-10-216-355-4
27	39	44.3	55	11	US-09-864-408A-2362
28	39	44.3	69	14	US-10-348-713-3
29	39	44.3	90	14	US-10-106-698-4634
30	39	44.3	115	14	US-10-156-761-9869
31	39	44.3	268	9	US-09-949-192-53
32	39	44.3	376	15	US-10-622-896-2
33	39	44.3	381	15	US-10-622-896-4
34	39	44.3	394	14	US-10-081-816-40
35	39	44.3	397	15	US-10-447-328-60
36	39	44.3	479	15	US-10-369-493-20600
37	39	44.3	1263	10	US-09-882-694-11
38	39	44.3	1548	15	US-10-369-493-6347
39	38.5	43.8	560	9	US-09-931-212A-4
40	38.5	43.8	560	9	US-09-915-181A-5
41	38.5	43.8	560	10	US-09-965-522-4
42	38.5	43.8	560	15	US-10-314-790-5
43	38.5	43.8	578	9	US-09-740-041-4
44	38.5	43.8	578	14	US-10-389-967-4
45	38.5	43.8	582	9	US-09-915-181A-4

ALIGNMENTS

RESULT 1

US-10-036-444-5
; Sequence 5, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; TITLE OF INVENTION: "No. US20020142445A1e1 triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 5
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-5

Query Match 100.0%; Score 88; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAVG 19
Db 1 VLLLRAGFYAVSFLSVAVG 19

RESULT 2

US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1

```

; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
; US-10-036-444-2

```

```

Query Match      100.0%; Score 88; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. NO. 7.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VLLLRAGFYAVSFLSVAVG 19
Db 139 VLLLRAGFYAVSFLSVAVG 157

```

```

RESULT 3
US-10-121-393-2
; Sequence 2, Application US/10121393
; Publication No. US20030115633A1
; GENERAL INFORMATION:
; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES RELATING TO PAPAYA RIPENING
; FILE REFERENCE: 19603/3001
; CURRENT APPLICATION NUMBER: US/10/121.393
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,008
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Carica papaya
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa at position 143 in this sequence is any amino
; OTHER INFORMATION: acid
; US-10-121-393-2

```

```

Query Match      54.5%; Score 48; DB 14; Length 665;
Best Local Similarity 62.5%; Pred. NO. 13;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
Db 383 LRAGINKISILSIAGV 398

```

```

RESULT 4
US-10-121-539-2
; Sequence 2, Application US/10121539
; Publication No. US20030204869A1
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje

```

```

; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Chiang, Chu-Hui
; TITLE OF INVENTION: A METHOD TO CONTROL THE RIPENING OF PAPAYA FRUIT AND
; TITLE OF INVENTION: CONFER DISEASE RESISTANCE TO PAPAYA PLANTS
; FILE REFERENCE: 19603/3601
; CURRENT APPLICATION NUMBER: US/10/121,539
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,022
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Carica papaya
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa at position 143 in this sequence is any amino
; OTHER INFORMATION: acid
; US-10-121-539-2

```

```

Query Match      54.5%; Score 48; DB 15; Length 665;
Best Local Similarity 62.5%; Pred. NO. 13;
Matches 10; Conservative 2; Mismatches 4; Indels 4; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
Db 383 LRAGINKISILSIAGV 398

```

```

RESULT 5
US-10-121-393-6
; Sequence 6, Application US/10121393
; Publication No. US20030115633A1
; GENERAL INFORMATION:
; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES RELATING TO PAPAYA RIPENING
; FILE REFERENCE: 19603/3001
; CURRENT APPLICATION NUMBER: US/10/121.393
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,008
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Carica papaya
; US-10-121-393-6

```

```

Query Match      54.5%; Score 48; DB 14; Length 721;
Best Local Similarity 68.8%; Pred. NO. 14;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
Db 526 LRAGVNVSLLSIAGV 541

```

```

RESULT 6
US-10-121-539-6
; Sequence 6, Application US/10121539
; Publication No. US20030204869A1
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Chiang, Chu-Hui
; TITLE OF INVENTION: A METHOD TO CONTROL THE RIPENING OF PAPAYA FRUIT AND
; TITLE OF INVENTION: CONFER DISEASE RESISTANCE TO PAPAYA PLANTS

```

; FILE REFERENCE: 19603/3601
; CURRENT APPLICATION NUMBER: US/10/121,539
; PRIOR FILING DATE: 2002-04-11
; PRIOR FILING DATE: 2002-04-11
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Carica papaya
US-10-121-539-6

Query Match 54.5%; Score 48; DB 15; Length 721;
Best Local Similarity 68.8%; Pred. No. 14;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 LRAGFYAVSFSLVAVG 19
Db 526 LRAGVNVKVSLSIAVG 541

RESULT 7
US-10-369-493-10510
; Sequence 10510, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10510
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Spingomonas aromaticivorans
US-10-369-493-10510

Query Match 50.6%; Score 44.5; DB 15; Length 473;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

Qy 3 LLRAGF---YAVSFSLVAVG 19
Db 95 IAKAGFETRFVAVSFIRAAIG 114

RESULT 8
US-09-847-102A-42
; Sequence 42, Application US/09847102A
; Publication No. US2003004409A1
; GENERAL INFORMATION:
; APPLICANT: University of California
; APPLICANT: Carson, Dennis A.
; APPLICANT: Carr, Maripat
; APPLICANT: Rhee, Chae-Seo
; APPLICANT: Lorenzo, Leoni M.
; APPLICANT: Malini, Sen
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS
; FILE REFERENCE: 22000-20629.00
; CURRENT APPLICATION NUMBER: US/09/847,102A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 42
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Drosophila
US-09-847-102A-42

Query Match 48.9%; Score 43; DB 10; Length 536;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LLLRAGFYAVSFSLVAVG 19
Db 424 LMLRIGFFSGILFILPAVG 441

RESULT 9
US-09-738-626-4562
; Sequence 4562, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4562
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4562

Query Match 48.9%; Score 43; DB 9; Length 578;
Best Local Similarity 52.6%; Pred. No. 74;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VLLLRAGFYAVSFSLVAVG 19
Db 265 VLILGTGFLVNLIDLVTVG 283

RESULT 10
US-09-740-041-2
; Sequence 2, Application US/09740041
; Patent No. US20020082190A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589

```

; TYPE: PRT
; ORGANISM: Human
US-09-740-041-2

Query Match          47.2%; Score 41.5; DB 9; Length 589;
Best Local Similarity 52.2%; Pred. No. 1.4e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy 2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db 403 LLLVVGFSHTKGVAISFLVAVG 425
    |||||

RESULT 11
US-10-389-967-2
; Sequence 2, Application US/10389967
; Publication No. US20030166153A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USE THEREOF
; FILE REFERENCE: CL001001-DIV
; CURRENT APPLICATION NUMBER: US/10/389,967
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-389-967-2

Query Match          47.2%; Score 41.5; DB 14; Length 589;
Best Local Similarity 52.2%; Pred. No. 1.4e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy 2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db 403 LLLVVGFSHTKGVAISFLVAVG 425
    |||||

RESULT 12
US-09-915-181A-3
; Sequence 3, Application US/09915181A
; Patent No. US20020098473A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAUX, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-932610US
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-3

Query Match          47.2%; Score 41.5; DB 9; Length 850;
Best Local Similarity 52.2%; Pred. No. 2e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy 2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db 416 LLLVVGFSHTKGVAISFLVAVG 438
    |||||

RESULT 13
US-09-738-626-6265
; Sequence 6265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6265
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6265

Query Match          46.6%; Score 41; DB 9; Length 297;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 VLLLRAGFYAVSFSLVAV 18
    |||||
Db 249 VILLELAIFALSFLGSAV 266
    |||||

RESULT 14
US-10-156-761-9631
; Sequence 9631, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9631
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9631

Query Match          46.6%; Score 41; DB 14; Length 320;
Best Local Similarity 57.1%; Pred. No. 84;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

Qy 2 LLLRAGFYAVSFLS 15
 Db 190 LNLQAGFYMIFLA 203

RESULT 15
 US-10-080-170-441
 ; Sequence 441, Application US/10080170
 ; Publication No. US20030129601A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COLE, S.T.
 ; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
 ; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
 ; TREATMENT OF MYCOBACTERIOSES
 ; FILE REFERENCE: 03495.0218
 ; CURRENT APPLICATION NUMBER: US/10/080,170
 ; CURRENT FILING DATE: 2002-06-10
 ; PRIOR APPLICATION NUMBER: 60/270,123
 ; PRIOR FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 652
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 441
 ; LENGTH: 333
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-10-080-170-441

Query Match 46.6%; Score 41; DB 14; Length 333;
 Best Local Similarity 55.6%; Pred. No. 88;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LLLRAGFYAVSFLS 19
 Db 46 VLLRQGEPAVSFLISSG 63

Search completed: February 26, 2004, 12:33:55
 Job time : 8.60212 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 3.67905 Seconds
(without alignments)
266.616 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	53.4	838	2	US-08-696-944-19
2	44	50.0	87	4	US-09-107-532A-7259
3	44	50.0	385	4	US-09-491-577-20
4	44	50.0	602	4	US-09-134-001C-4416
5	42	47.7	435	4	US-09-252-991A-17750
6	41.5	47.2	589	4	US-09-740-041-2
7	40	45.5	263	4	US-09-134-000C-5703
8	40	45.5	272	4	US-09-252-991A-24088
9	40	45.5	385	4	US-09-328-352-5641
10	40	45.5	394	4	US-08-543-681A-5779
11	39	44.3	69	3	US-08-783-974-3
12	39	44.3	69	4	US-09-546-306-3
13	39	44.3	726	4	US-09-252-991A-28519
14	39	44.3	730	2	US-08-696-944-2
15	39	44.3	740	4	US-09-489-039A-11157
16	39	44.3	1263	4	US-09-351-224E-11
17	39	44.3	1263	4	US-09-677-488A-11
18	39	44.3	1263	4	US-09-677-682B-11
19	38.5	43.8	198	4	US-09-328-352-8198
20	38.5	43.8	560	1	US-08-647-484-2
21	38.5	43.8	560	1	US-08-647-481-2
22	38.5	43.8	560	1	US-08-430-033A-2
23	38.5	43.8	560	2	US-08-805-118-4
24	38.5	43.8	560	4	US-09-391-958-4
25	38.5	43.8	560	5	PCT-US96-05792-2
26	38.5	43.8	578	4	US-09-740-041-4
27	38	43.2	236	3	US-09-121-979-4

28	38	43.2	236	3	US-09-332-319-4	Sequence 4, Appli
29	38	43.2	236	4	US-09-239-857-2	Sequence 2, Appli
30	38	43.2	493	4	US-09-543-681A-4617	Sequence 4617, Ap
31	38	43.2	555	4	US-09-543-681A-4582	Sequence 4582, Ap
32	38	43.2	575	4	US-09-489-039A-10911	Sequence 10911, A
33	38	43.2	739	1	US-07-618-946B-22	Sequence 22, Appl
34	38	43.2	814	1	US-07-618-946B-23	Sequence 23, Appl
35	37.5	42.6	760	4	US-09-252-991A-18711	Sequence 18711, A
36	37	42.0	186	4	US-09-134-000C-4177	Sequence 4177, Ap
37	37	42.0	238	4	US-09-134-000C-3467	Sequence 3467, Ap
38	37	42.0	256	4	US-09-489-039A-11447	Sequence 11447, A
39	37	42.0	263	4	US-09-489-039A-13979	Sequence 13979, A
40	37	42.0	298	2	US-08-838-543-5	Sequence 5, Appli
41	37	42.0	336	1	US-08-118-270-50	Sequence 50, Appl
42	37	42.0	336	5	PCT-US93-08528-50	Sequence 50, Appl
43	37	42.0	349	4	US-09-489-039A-11026	Sequence 11026, A
44	37	42.0	376	1	US-08-614-801A-6	Sequence 6, Appli
45	37	42.0	404	4	US-08-630-915A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-08-696-944-19

; Sequence 19, Application US/08696944

; Patent No. 5981831

; GENERAL INFORMATION:

; APPLICANT: Sumant CHENGAPPA

; APPLICANT: Susan A. HELLYER

; APPLICANT: John S. REID

; APPLICANT: Jacqueline DE SILVA

; TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.

; STREET: 1100 New York Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MS Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/696,944

; FILING DATE: 23-AUG-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB95/00372

; FILING DATE: 23-FEB-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9403423.8

; FILING DATE: 23-FEB-1994

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 838 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-696-944-19

Query Match 53.4%; Score 47; DB 2; Length 838;

Best Local Similarity 62.5%; Pred No. 9;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19

Db 530 LRAGVKNISLLSIAGV 545

RESULT 2

US-09-107-532A-7259
; Sequence 7259, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A. Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/985,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7259:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...87
; SEQUENCE DESCRIPTION: SEQ ID NO: 7259:
US-09-107-532A-7259
Query Match 50.0%; Score 44; DB 4; Length 87;
Best Local Similarity 60.0%; Pred. No. 2.2;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFLS 15
||:|||||::||
DB 15 VLMIRAGFPATITLS 29
RESULT 3
US-09-491-577-20
; Sequence 20, Application US/09491577
; Patent No. 6610511
; GENERAL INFORMATION:
; APPLICANT: Yale University
; APPLICANT: Carlson, John R.
; APPLICANT: Kim, Huhhyong
; APPLICANT: Clyne, Peter J.
; APPLICANT: Warr, Coral G.
; TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila
; FILE REFERENCE: 44574-5061-US
; CURRENT APPLICATION NUMBER: US/09/491,577
; CURRENT FILING DATE: 2000-01-25

; EARLIER APPLICATION NUMBER: US 60/117,132
; EARLIER FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-491-577-20
Query Match 50.0%; Score 44; DB 4; Length 385;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 VLLLRAGFYAVSFLS 15
||:|||||::||
DB 264 VLVLTANFYAVIAVLS 278
RESULT 4
US-09-134-001C-4416
; Sequence 4416, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4416
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4416
Query Match 50.0%; Score 44; DB 4; Length 602;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 7 GFVAVSFLSVAV 18
||||:||||:
DB 272 GFVAVSFLFLV 283
RESULT 5
US-09-252-991A-17750
; Sequence 17750, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17750
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17750
Query Match 47.7%; Score 42; DB 4; Length 435;

Thu Feb 26 12:38:26 2004

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24088
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24088

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 272;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 151 VVYSLQAVGFLSVAG 166

RESULT 9
US-09-328-352-5641
; Sequence 5641, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5641
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5641

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 385;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLV 16
DB 14 LLVLSGAFYSIDFLKV 29

RESULT 10
US-09-543-681A-5779
; Sequence 5779, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5779
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-5779

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 394;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 GFYAVSFSLVAVG 19
DB 223 GWYIVGFLAIG 235

RESULT 8
US-09-252-991A-24088
; Sequence 24088, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24088
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-24088

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 272;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 151 VVYSLQAVGFLSVAG 166

RESULT 9
US-09-328-352-5641
; Sequence 5641, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5641
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5641

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 385;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLV 16
DB 14 LLVLSGAFYSIDFLKV 29

RESULT 10
US-09-543-681A-5779
; Sequence 5779, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5779
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-5779

Query Match
Best Local Similarity 45.5%; Score 40; DB 4; Length 394;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 GFYAVSFSLVAVG 19
DB 223 GWYIVGFLAIG 235

RESULT 8
US-09-252-991A-24088
; Sequence 24088, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

```



```
Best Local Similarity 69.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFL 14
Db 274 LLLRGHLAGVFL 286

RESULT 14
US-08-696-944-2
; Sequence 2, Application US/08696944
; Patent No. 5981831
; GENERAL INFORMATION:
; APPLICANT: Sumant CHENGAPPA
; APPLICANT: Susan A. HELLYER
; APPLICANT: John S. REID
; APPLICANT: Jacqueline DE SILVA
; TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,944
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00372
; FILING DATE: 23-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9403423.8
; FILING DATE: 23-FEB-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-696-944-2

Query Match 44.3%; Score 39; DB 2; Length 730;
Best Local Similarity 56.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAG 19
Db 536 LRVGNNKISLLSVSG 551

RESULT 15
US-09-489-039A-11157
; Sequence 11157, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11157
; LENGTH: 740
```

```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11157

Query Match 44.3%; Score 39; DB 4; Length 740;
Best Local Similarity 60.0%; Pred. No. 1.8e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSV 16
Db 654 LLIRCGFKANSFAGV 668

Search completed: February 26, 2004, 12:16:17
Job time : 5.67905 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 4.98939 Seconds
(without alignments)
636.214 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.78.*

2: PIR1.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	60	29.4	1801	1	WMRTS
2	59.5	28.2	660	1	QBEB3
3	59	28.9	544	2	T36645
4	58.5	28.7	116	2	T46473
5	58	28.4	180	2	E84768
6	57	27.9	435	2	S40993
7	56.5	27.7	1321	2	J80352
8	55	27.0	1620	2	T27283
9	54.5	26.7	147	2	S37485
10	54.5	26.7	3078	2	T28432
11	54	26.5	1300	2	A36502
12	53.5	26.2	2767	1	U1HU
13	53	26.0	460	2	D97679
14	53	26.0	597	2	T16006
15	52.5	25.7	91	2	S37486
16	52.5	25.7	153	2	PNQ564
17	52.5	25.7	577	2	B37057
18	52.5	25.7	799	2	A38308
19	52	25.5	137	2	A87586
20	52	25.5	1798	2	S33869
21	51.5	25.2	106	2	E95330
22	51.5	25.2	138	2	T48808
23	51.5	25.2	378	2	B59180
24	51.5	25.2	425	2	T48724
25	51.5	25.2	436	2	D88826
26	51.5	25.2	438	2	S28263
27	51.5	25.2	929	2	T51932
28	51.5	25.2	1166	1	S06142
29	51.5	25.2	1344	2	S47412

ALIGNMENTS

RESULT 1

WMRTS

Laminin beta-2 chain precursor - rat

N/Alternate names: laminin chain B3; S-laminin

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-Dec-1999

C/Accession: S03539

R/Hunter, D.D.; Shah, V.; Merlie, J.P.; Sanes, J.R.

Nature 338, 229-234, 1989

A/Title: A laminin-like adhesive protein concentrated in the synaptic cleft of the neur

A/Reference number: S03539; MUID:89159410; PMID:2922051

A/Accession: S03539

A/Molecule type: mRNA

A/Residues: 1-1801 <HUN>

A/Cross-references: EMBL:X16563; NID:G57250; PIDN:CAA34561.1; PID:G57251

C/Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin

C/Function:

A/Description: interact with cells and with other basement membrane proteins to promote

C/Superfamily: laminin beta-1 chain; laminin-type EGF-like homology

C/Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular

F/1-35/Domain: signal sequence #status predicted <SIG>

F/36-1801/Product: laminin beta-2 chain #status predicted <MAT>

F/36-285/Domain: VI <DOM6>

F/286-555/Domain: V <DOM5>

F/286-347/Domain: laminin-type EGF-like homology <LE01>

F/350-410/Domain: laminin-type EGF-like homology <LE02>

F/413-470/Domain: laminin-type EGF-like homology <LE03>

F/473-522/Domain: laminin-type EGF-like homology <LE04>

F/525-555/Domain: laminin-type EGF-like homology #status atypical <LE05>

F/556-784/Domain: IV <DOM4>

F/786-831/Domain: laminin-type EGF-like homology <LE06>

F/788-1196/Domain: III <DOM3>

F/834-877/Domain: laminin-type EGF-like homology <LE07>

F/880-927/Domain: laminin-type EGF-like homology <LE08>

F/930-986/Domain: laminin-type EGF-like homology <LE09>

F/989-1038/Domain: laminin-type EGF-like homology <LE10>

F/1041-1095/Domain: laminin-type EGF-like homology <LE11>

F/1098-1143/Domain: laminin-type EGF-like homology <LE12>

F/1146-1190/Domain: laminin-type EGF-like homology <LE13>

F/1197-1412/Domain: II <DOM2>

F/1197-1412/Region: heptad repeats

F/1413-1445/Domain: alpha <ALP>

F/1446-1801/Region: heptad repeats

F/1446-1801/Domain: I <DOM1>

F/45-50/Disulfide bonds: #status predicted

F/251-371.1088.1252.1311.1351.1502/Binding site: carbohydrate (Asn) (covalent) #status

F/1193.1196.1800/Disulfide bonds: interchain #status predicted

Query Match 29.4%; Score 60; DB 1; Length 1801;

Best Local Similarity 29.8%; Pred. No. 21;

Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

QY 5 YQKCHCHM-----GTHCHSD-GPRGVIPERPC 32
 Db 1113 FTGQCHCHAGRGRTCCQELHWGDFGLQCRACDCDPRG-IDKPCQ 1158

RESULT 2

QBE3
 HBLF1 protein - human herpesvirus 4 (strain B95-8)
 C/Species: human herpesvirus 4, Epstein-Barr virus
 C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 23-Aug-1997
 C/Accession: A03742
 R/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983
 A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
 A/Reference number: A93065; MUID:85035713; PMID:6092825
 A/Accession: A03742
 A/Molecule type: DNA
 A/Residues: 1-660 <BAN>
 R/Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
 Nature 310, 207-211, 1984
 A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A/Reference number: A03794; MUID:84270667; PMID:6087149
 A/Contents: annotation; protein coding region
 C/Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-52
 C/Superfamily: human herpesvirus 4 BHV1 protein

Query Match 29.2%; Score 59.5; DB 1; Length 660;
 Best Local Similarity 57.1%; Pred. No. 10;
 Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 13 MGTCHSSDGRGVIPERPC 33
 Db 1 MGTQCQARGPR-TTLPHP 20

RESULT 3

T36645
 probable large integral membrane protein - Streptomyces coelicolor
 C/Species: Streptomyces coelicolor
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T36645
 R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1999
 A/Reference number: Z21610
 A/Accession: T36645
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-544 <OLI>
 A/Cross-references: EMBL:AL078610; PIDN:CAB44413.1; GSPDB:GN00070; SCOEDB:SCH35.14c
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SCOEDB:SCH35.14c

Query Match 28.9%; Score 59; DB 2; Length 544;
 Best Local Similarity 52.6%; Pred. No. 9.9;
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 12 HMGTHCHSSDGRGVIPERPC 30
 Db 434 HFGTHVLGDGPPAVVPH 452

RESULT 4

T46473
 hypothetical protein DKFZp3401230.1 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C/Accession: T46473
 R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A/Reference number: Z23034
 A/Accession: T46473
 A/Status: preliminary

Query Match 27.9%; Score 57; DB 2; Length 435;
 Best Local Similarity 38.7%; Pred. No. 14;
 Matches 12; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 2 TVTYQKCHCHMGTHCHSSDGRGVIPERPC 32
 Db 324 TCSLSGKC-CKLNHC-----PDGTPETSC 348

RESULT 7

JEO352

A/Molecule type: mRNA
 A/Residues: 1-116 <AAA>
 A/Cross-references: EMBL:AL137489
 A/Experimental source: adult testis; clone DKFZp3401230
 C/Genetics:
 A/Note: DKFZp3401230.1

Query Match 28.7%; Score 58.5; DB 2; Length 116;
 Best Local Similarity 43.5%; Pred. No. 2.9;
 Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 10 HCMGTHCHSSDGRGVIPERPC 32
 Db 51 HCHGLYRCHGRPQREGL---PRC 70

RESULT 5

E84768
 hypothetical protein At2g35430 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C/Accession: E84768
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: E84768
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-180 <STO>
 A/Cross-references: GB:AE002093; NID:g3608145; PIDN:AAC36178.1; GSPDB:GN00139
 C/Genetics:
 A/Gene: At2g35430
 A/Map position: 2

Query Match 28.4%; Score 58; DB 2; Length 180;
 Best Local Similarity 42.9%; Pred. No. 4.9;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 5 YQKCHCHMGTHCHSSDGRG 25
 Db 152 WQTTGYCPFGSHCHFAHPSG 172

RESULT 6

S40993
 hypothetical protein K04H4.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 03-May-1994 #sequence_revision 02-Aug-1994 #text_change 02-Aug-1994
 C/Accession: S40993
 R/Ainscough, R.
 submitted to the EMBL Data Library, October 1993
 A/Reference number: S40991
 A/Accession: S40993
 A/Molecule type: DNA
 A/Residues: 1-435 <AIN>
 A/Cross-references: EMBL:Z27078
 C/Genetics:
 A/Introns: 90/1; 305/3; 379/1; 425/1

mucin MUC5B, tracheobronchial - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
 C:Accession: J03352
 R:Officer, G.D.; Nunes, D.P.; Keates, A.C.; Afidhal, N.H.; Troxler, R.F.
 Biochem. Biophys. Res. Commun. 251, 350-355, 1998
 A:Title: The amino-terminal sequence of MUC5B contains conserved multifunctional D domain
 A:Reference number: J03352; MUID:99009274; PMID:9790959
 A:Accession: J03352
 A:Molecule type: mRNA
 A:Residues: 1-1321 <OFF>
 A:CROSS-references: GB:A086604; NID:g3789926; PIDN:AAC67545.1; PID:g3789927
 C:Comment: This protein is large multimeric glycoproteins which is secreted by epithelia
 C:Genetics:
 A:Gene: MUC5B
 Query Match 27.7%; Score 56.5; DB 2; Length 1321;
 Best Local Similarity 39.0%; Pred. No. 44;
 Matches 16; Conservative 0; Mismatches 10; Indels 15; Gaps 3;
 QY 7 GKCH-----CHMGTHCHS-----SDGPRGVIPERPCP 33
 DB 816 GSCHTLDVGC-FSTHCVSVCVCPPLVSDGSGGCIABEDCP 855
 RESULT 8
 T27283
 hypothetical protein Y64G10A.f - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27283
 R:Ainscough, R.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z20336
 A:Accession: T27283
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1620 <WIL>
 A:CROSS-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CBSP:Y64G10A.f
 A:Experimental source: clone Y64G10A
 C:Genetics:
 A:Gene: CBSP:Y64G10A.f
 A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7
 Query Match 27.0%; Score 55; DB 2; Length 1620;
 Best Local Similarity 46.2%; Pred. No. 80;
 Matches 12; Conservative 0; Mismatches 8; Indels 6; Gaps 1;
 QY 8 KCHCHMGTHCHSSDGPGRGVIPERPCP 33
 DB 1012 KDCADGHCDFSDG-----ECICP 1031
 RESULT 9
 S37485
 gene msg1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C:Accession: I48669; S37485
 R:Tronik-Le Roux, D.; Sencorale-Pose, M.; Rougeon, F.
 Gene 142, 175-182, 1994
 A:Title: Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice
 A:Reference number: I48669; MUID:94252564; PMID:8194749
 A:Accession: I48669
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-147 <RES>
 A:CROSS-references: EMBL:X71629; NID:g406256; PIDN:CAA50636.1; PID:g406257
 C:Genetics:
 A:Gene: msg1
 C:Superfamily: proline-rich peptide P-B
 Query Match 26.7%; Score 54.5; DB 2; Length 147;

Best Local Similarity 44.0%; Pred. No. 11;
 Matches 11; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
 QY 9 CHCHMGTHCHSSDGPGRGVIPERPCP 33
 DB 20 CECHRGPRHD---PRGPPPPPPP 41
 RESULT 10
 T28432
 variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N:Alternate names: erythrocyte membrane binding protein 1 (EMBL)
 C:Species: Plasmodium falciparum
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T28432
 R:Su, X.Z.; Heatwole, V.M.; Wertheimer, S.P.; Guinet, F.; Herrfeldt, J.A.; Peterson, D.
 Cell 82, 89-100, 1995
 A:Title: The large diverse gene family var encodes proteins involved in cytoadherence a
 A:Reference number: Z20487; MUID:95330813; PMID:7608788
 A:Accession: T28432
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-3078 <SUX>
 A:CROSS-references: EMBL:L40608; NID:g886374; PID:g886375; PIDN:AAA75396.1
 C:Genetics:
 A:Gene: var-1
 A:Introns: 2611/3
 Query Match 26.7%; Score 54.5; DB 2; Length 3078;
 Best Local Similarity 40.0%; Pred. No. 1.6e+02;
 Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;
 QY 5 YQKCHCHMGTHCHS-SDGPRGVIPERPCP 33
 DB 1698 YPEKCDYQGRKVPSPPPPPVQPPAP 1727
 RESULT 11
 A36502
 insulin receptor-related receptor precursor - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C>Date: 01-Feb-1991 #sequence_revision 01-Feb-1991 #text_change 23-May-1997
 C:Accession: A36502
 R:Shier, P.; Watt, V.M.
 J. Biol. Chem. 264, 14605-14608, 1989
 A:Title: Primary structure of a putative receptor for a ligand of the insulin family.
 A:Reference number: A36502; MUID:89359245; PMID:2768234
 A:Accession: A36502
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1300 <SHI>
 A:CROSS-references: GB:J05047
 C:Superfamily: insulin receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphoprotein; receptor; transmembrane protein
 F:977-1253/Domain: protein kinase homology <KIN>
 F:985-993/Region: protein kinase ATP-binding motif
 Query Match 26.5%; Score 54; DB 2; Length 1300;
 Best Local Similarity 52.6%; Pred. No. 87;
 Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;
 QY 4 YYQKCH-CHMGTHCHS 20
 DB 253 YYQSACHRACPLGTYESHES 271
 RESULT 12
 UIHU
 thyroglobulin precursor, major splice form - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Feb-1986 #sequence_revision 05-Nov-1999 #text_change 16-Jun-2000
 C:Accession: A59110; S00014; A01532; S03422; I38343; I57669; S02266; S39432; S66241; S6
 R:Malthiery, Y.

submitted to the EMBL Data Library, April 1988

A:Description: Human mRNA for thyroglobulin.
A:Reference number: A59110
A:Accession: A59110
A:Molecule type: mRNA
A:Status: translated from GB/EMBL/DBJ
A:Residues: 1-2767 <MAL1>
A:Cross-references: GB:X05615; NID:G37173; PIDN:CAA29104.1; PID:G37174
A:Note: revision to S00014
R:Malthery, Y.; Lissitzky, S.
Eur. J. Biochem. 165, 491-498, 1987
A:Title: Primary structure of human thyroglobulin deduced from the sequence of its 8448-
A:Reference number: S00014; MUID:8746630; PMID:3595599
A:Accession: S00014
A:Molecule type: mRNA
A:Status: translated from GB/EMBL/DBJ
A:Residues: 1-1041, 'Y', 1043-1057, 'T', 1059-2767 <MAL2>
A:Cross-references: GB:X05615; NID:G37173
A:Note: this sequence is revised in A59110
R:Malthery, Y.; Lissitzky, S.
Eur. J. Biochem. 147, 53-58, 1985
A:Title: Sequence of the 5'-end quarter of the human-thyroglobulin messenger ribonucleic
A:Reference number: A01532; MUID:85127024; PMID:3971976
A:Accession: A01532
A:Molecule type: mRNA
A:Status: translated from GB/EMBL/DBJ
A:Residues: 1-730 <MAW>
A:Cross-references: GB:X02154; NID:G37175; PIDN:CAA26089.1; PID:gl353349
A:Note: the translated sequence in GenBank entry HSTHYR5, release 111.0, (PIDN:CAA26089
n 1-Mat
R:Parma, J.; Christophe, D.; Pohl, V.; Vassart, G.
J. Mol. Biol. 196, 769-779, 1987
A:Title: Structural organization of the 5' region of the thyroglobulin gene. Evidence for
A:Reference number: S03422; MUID:88062712; PMID:3681978
A:Accession: S03422
A:Molecule type: DNA
A:Residues: 1-134, 'Q', 136-415, 640-652, 'G', 654-733, 'A', 735-737, 880-983, 'DR', 985-999 <PAR>
A:Cross-references: EMBL:X06059; NID:G37145; PIDN:CAA29454.1; PID:gl359884; EMBL:X06067;
R:Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G.
Nucleic Acids Res. 13, 5127-5144, 1985
A:Title: An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream of
A:Reference number: I38344; MUID:85269632; PMID:2991855
A:Accession: I38343
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22, 'GRF', <CHR>
A:Cross-references: EMBL:X02749; NID:G37162; PIDN:CAA26527.1; PID:g758106
R:Targovnik, H.M.; Cochaux, P.; Corach, D.; Vassart, G.
Mol. Cell. Endocrinol. 84, R23-R26, 1992
A:Title: Identification of a minor Ig mRNA transcript in RNA from normal and goitrous th
A:Reference number: I57669; MUID:92347597; PMID:1639210
A:Accession: I57669
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1503-1508, 'L', 1567-1601 <REW>
A:Cross-references: GB:S40807; NID:G252170; PIDN:AA822685.1; PID:G252171
A:Note: this sequence fragment represents a minor splice form
R:Marriq, C.; Lejeune, P.J.; Venot, N.; Vinet, L.
FEBS Lett. 242, 414-418, 1989
A:Title: Hormone synthesis in human thyroglobulin: possible cleavage of the polypeptide
A:Reference number: S02266; MUID:8912111; PMID:2914619
A:Contents: evidence for Tyr-149 as ring donor for thyroxine formation
A:Accession: S02266
A:Molecule type: protein
A:Residues: 101-109; 114-121; 126-131; 143-149 <MAR>
A:Note: only the first peptide was sequenced; others were isolated and their amino acid
R:Gentile, F.; Salvatore, G.
Eur. J. Biochem. 218, 603-621, 1993
A:Title: Preferential sites of proteolytic cleavage of bovine, human and rat thyroglobu
A:Reference number: S39431; MUID:94094855; PMID:8269951
A:Accession: S39432
A:Molecule type: Protein
A:Residues: 20-27; 522-527; 541-547; 616-619, 'X', 621; 1001-1005; 1009-1011, 'X', 1013; 1424-1430
R:Xiao, S.; Pollock, H.G.; Taurag, A.; Rawitch, A.B.
Arch. Biochem. Biophys. 320, 96-105, 1995

A:Title: Characterization of hormonogenic sites in an N-terminal, cyanogen bromide frag
A:Reference number: S66241; MUID:95314327; PMID:7793989
A:Accession: S66241
A:Molecule type: protein
A:Residues: 'D', 21-23; 45-47; 147-148; 150-152 <XIA>
R:Yang, S.X.; Pollock, H.G.; Rawitch, A.B.
Arch. Biochem. Biophys. 327, 61-70, 1996
A:Title: Glycosylation in human thyroglobulin: location of the N-linked oligosaccharide
A:Reference number: S62778; MUID:96201348; PMID:8615697
A:Accession: S62778
A:Molecule type: protein
A:Residues: 69-84; 136-210, 'T', 212-213; 476-492; 523-539; 741-770; 811-848; 850-853; 938-950; 1
6-1783, 'D', 1785, 'R', 1999-2017; 2241-2258; 2270-3285, 'P', 2287-2307; 2578-2584 <YAN>
R:Ieiri, T.; Cochaux, P.; Targovnik, H.M.; Suzuki, M.; Shimoda, S.; Perret, J.; Vassart
J. Clin. Invest. 88, 1901-1905, 1991
A:Title: A 3' splice site mutation in the thyroglobulin gene responsible for congenital
A:Reference number: I55565; MUID:92091498; PMID:1752952
A:Accession: I55565
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 90-91, 160-161 <IEI>
A:Cross-references: GB:S71821; NID:G240935; PIDN:AA820665.1; PID:G240936
A:Note: mutant splice form
C:Comment: The thyroglobulin molecule is produced in the thyroid gland and is the precu
C:Genetics:

A:Gene: GDB:TG
A:Cross-references: GDB:I20434; OMIM:188450
A:Map position: 8q24.2-8q24.3
A:Introns: 23/1; 59/2; 92/1; 160/1; 213/2; 249/1; 297/1; 359/1; 726/1; 921/1; 1000/1 #8
C:Complex: homodimer
C:Function:
A:Description: precursor of thyroid hormones thyroxine and triiodothyronine; iodine stc
C:Superfamily: thyroglobulin; cholinesterase homology; glycoprotein; homodimer; iodine; thyroid
C:Keywords: alternative splicing; duplication; glycoprotein; homodimer; iodine; thyroid
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2767/Product: thyroglobulin #status predicted <MAT>
F:34-92/Domain: thyroglobulin type I repeat homology <THV1>
F:96-160/Domain: thyroglobulin type I repeat homology <THY2>
F:164-297/Domain: thyroglobulin type I repeat homology #status atypical <THV3>
F:301-358/Domain: thyroglobulin type I repeat homology <THY4>
F:608-658/Domain: thyroglobulin type I repeat homology <THY5>
F:662-726/Domain: thyroglobulin type I repeat homology <THY6>
F:730-921/Domain: thyroglobulin type I repeat homology #status atypical <THY7>
F:925-1072/Domain: thyroglobulin type I repeat homology #status atypical <THY8>
F:1076-1144/Domain: thyroglobulin type I repeat homology <THY9>
F:1148-1209/Domain: thyroglobulin type I repeat homology <TH10>
F:1455-1468/Region: type II repeat
F:1469-1485/Region: type II repeat
F:1486-1502/Region: type II repeat
F:1602-1722/Region: type IIIa repeat
F:1723-1891/Region: type IIIb repeat
F:1892-1994/Region: type IIIa repeat
F:1995-2128/Region: type IIIb repeat
F:2129-2186/Region: type IIIa repeat
F:2227-2725/Domain: cholinesterase homology <ACE>
F:24, 2572, 2586/Modified site: thyroxine (Tyr) #status predicted
F:76, 198, 484, 529, 748, 816, 947, 1219, 1348, 1364, 1715, 1773, 2012, 2249, 2294, 2581/Binding site:
F:110, 496, 1868, 2121/Binding site: carbohydrate (Asn) (covalent) #status absent
F:149/Modified site: dehydralanine (Tyr) #status predicted
F:2765/Modified site: triiodothyronine (Tyr) #status predicted

Query Match 26.2%; Score 53.5; DB 1; Length 2767;
Best Local Similarity 31.2%; Pred. No. 1.9e-02;
Matches 15; Conservative 2; Mismatches 12; Indels 19; Gaps 3;

QY 5 YQKGC-----HCHMGT-HCHSSDGRGVPE-----PRCP 33
DB 1026 YMPQDAFGSWPVCAGTGHWCVDKGGFPGSLTARSLOIPQCP 1073

RESULT 13
D97679
argininosuccinate synthase (PA3525) [imported] - Agrobacterium tumefaciens (strain C58,

C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: D97679
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: D97679
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-460 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK88389.1; PID:gi5157878; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 4836
 A:Map position: circular chromosome
 C:Superfamily: argininosuccinate synthase

Query Match 26.0%; Score 53; DB 2; Length 460;
 Best Local Similarity 40.0%; Pred. No. 46;
 Matches 12; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 4 YVQKCHCHGTHCHSSDGRGVPEPRCP 33
 DB 18 YLTKRFTFGLHCKSAEIPR----KWRCP 43

RESULT 14

T16006
 protein kinase C3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 C:Accession: T16006; T37253
 R:Chisoe, S.
 submitted to the EMBL Data Library, September 1995
 A:Description: The sequence of C. elegans cosmid F09E5.
 A:Reference number: Z18444
 A:Accession: T16006
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-597 <CHI>
 A:Cross-references: EMBL:U37429; NID:gi1019949; PID:gi1019950; PIDN:AAA79341.1; CESP:F09E5
 A:Experimental source: strain Bristol N2
 R:Wu, S.L.; Rubin, C.S.
 submitted to the EMBL Data Library, September 1997
 A:Description: Molecular characterization of an atypical protein kinase C from C. elegans
 A:Reference number: Z21647
 A:Accession: T37253
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-597 <WUS>
 A:Cross-references: EMBL:AF025666; PIDN:AAB88885.1
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: F09E5.1; pkc-3
 A:Map position: 11
 A:Introns: 22/3; 109/1; 137/2; 179/3; 322/3; 442/2; 556/3
 C:Superfamily: protein kinase C zeta; protein kinase C zinc-binding repeat homology; protein kinase C zinc-binding repeat homology <KZN>
 F:251-522/Domain: protein kinase homology <KIN>

Query Match 26.0%; Score 53; DB 2; Length 597;
 Best Local Similarity 50.0%; Pred. No. 58;
 Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 2;

QY 8 KCHCHGTHC-HSSDGRGVIP 28
 DB 168 KCHRVTHCGCALQSP-NIIP 188

RESULT 15

S37486
 gene msg3 protein - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
 C:Accession: I48671; S37486
 R:Tronik-Le Roux, D.; Senorale-Pose, M.; Rougeon, P.
 Gene 142, 175-182, 1994
 A:Title: Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice
 A:Reference number: I48669; MUID:94252564; PMID:8194749
 A:Accession: I48671
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-91 <RES>
 A:Cross-references: EMBL:X71631; NID:g406260; PIDN:CRA50638.1; PID:g406261
 C:Genetics:
 A:Gene: msg3
 C:Superfamily: proline-rich peptide P-B

Query Match 25.7%; Score 52.5; DB 2; Length 91;
 Best Local Similarity 45.5%; Pred. No. 13;
 Matches 10; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSSDGRGVPEP 30
 DB 20 CECHRGPRRD---PRGPPFP 38

Search completed: February 26, 2004, 12:14:56
 Job time : 7.98939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 2.80106 Seconds
(without alignments)
613.452 Million cell updates/sec

Title: US-10-036-444-6
Perfect score: 204
Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	29.4	1799	1	LMB2 MOUSE
2	60	29.4	1801	1	LMB2 RAT
3	59.5	29.2	660	1	YHL1 EBV
4	58.5	28.7	2471	1	NTC2 MOUSE
5	54.5	26.7	147	1	SMR1 MOUSE
6	54	26.5	1300	1	TRR_CAVPO
7	53.5	26.2	2768	1	THYG_HUMAN
8	52.5	25.7	577	1	ITB6_CAVPO
9	52.5	25.7	589	1	ZN44_HUMAN
10	52.5	25.7	799	1	ITB5_HUMAN
11	52.5	25.7	1565	1	DMN_HUMAN
12	52.5	25.7	2813	1	VWF_CANFA
13	52	25.5	1798	1	LMB2_HUMAN
14	51.5	25.2	147	1	HMR1_RHIME
15	51.5	25.2	378	1	WIF1_BRARE
16	51.5	25.2	438	1	LIN3_CABEL
17	51.5	25.2	461	1	Y514_HUMAN
18	51.5	25.2	1167	1	XMRK_XIPWA
19	51.5	25.2	5703	1	MUSB_HUMAN
20	51	25.0	178	1	NEL1_MOUSE
21	51	25.0	325	1	VT2_SFVKA
22	51	25.0	398	1	DH2_XENLA
23	51	25.0	434	1	UROK_CHICK
24	51	25.0	699	1	ECM2_HUMAN
25	51	25.0	2482	1	VNF_PIG
26	50.5	24.8	74	1	EDF2_HUMAN
27	50.5	24.8	106	1	Y402_METJA
28	50.5	24.8	515	1	K34L_CABEL
29	50.5	24.8	655	1	ITB5_PAPCY
30	50	24.5	396	1	DH11_XENLA
31	50	24.5	759	1	MTS1_MOUSE
32	50	24.5	5376	1	ZAN_MOUSE
33	49	24.0	176	1	TR23_MOUSE

Q9h293 homo sapien
P77365 escherichia
P29825 myxoma viru
O63371 rattus norv
O14513 homo sapien
P28797 cavia porce
Q9qy70 z adam 15 p
P78325 homo sapien
O43052 schizosacch
P58512 homo sapien
P25117 sus scrofa
P48467 neurospora

34 49 24.0 177 1 I17E HUMAN
35 49 24.0 285 1 YAFY ECOLI
36 49 24.0 326 1 VT2_MXVL
37 49 24.0 328 1 P2Y5 RAT
38 49 24.0 416 1 NAPS_HUMAN
39 49 24.0 591 1 GRN_CAVPO
40 49 24.0 816 1 AD15 RAT
41 49 24.0 824 1 AD08 HUMAN
42 49 24.0 1150 1 RGAL SCHPO
43 48.5 23.8 204 1 CU67 HUMAN
44 48.5 23.8 498 1 CALR_PIG
45 48.5 23.8 928 1 KINH_NEUCR

ALIGNMENTS

RESULT 1
LMB2_MOUSE
ID LMB2_MOUSE STANDARD; PRT; 1799 AA.
AC Q61292; Q62182;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Laminin beta-2 chain precursor (S-laminin) (S-LAM).
GN LAMB2 OR LAMS
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/J;
RX MEDLINE=96278760; PubMed=8662701;
RA Durkin M.E., Gautam M., Loechel S., Sanes J.R., Merlie J.P.,
RA Albrechtsen R., Wewer U.M.;
RT "Structural organization of the human and mouse laminin beta2 chain
RT genes, and alternative splicing at the 5' end of the human
RT transcript.";
RL J. Biol. Chem. 271:13407-13416(1996).
RN [2]
RP SEQUENCE OF 348-428 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94319092; PubMed=8043959;
RA Aberdam D., Galliano M.F., Mattei M.-G., Ortonne J.P., Meneguzzi G.;
RT "S-laminin gene (Lams) maps to F1 band of mouse chromosome 9.";
RL Mamm. Genome 5:393-394(1994).
RN [3]
RP FUNCTION.
RC STRAIN=129/J;
RX MEDLINE=95191650; PubMed=7885444;
RA Noakes P.G., Gautam M., Mudd J., Sanes J.R., Merlie J.P.;
RT "Aberrant differentiation of neuromuscular junctions in mice lacking
RT s-laminin/laminin beta 2.";
RL Nature 374:258-262(1995).
CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration and organization
CC of cells into tissues during embryonic development By interacting
CC with other extracellular matrix components.
CC -!- FUNCTION: Laminin-3 (S-laminin) regulates the formation of motor
CC nerve terminals.
CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end. The beta-2 chain is a subunit of laminin-3 (S-laminin),
CC laminin-4 (S-merosin), and laminin-7 (KS-laminin).
CC -!- SUBCELLULAR LOCATION: Extracellular; found in the basement
CC membranes (major component).
CC -!- TISSUE SPECIFICITY: Neuromuscular synapse and kidney glomerulus.
CC -!- DOMAIN: The alpha-helical domains I and II are thought to interact
CC with other laminin chains to form a coiled coil structure.
CC -!- DOMAIN: Domains VI and IV are globular.

FT	DISULFID	832	844	BY SIMILARITY.
FT	DISULFID	834	851	BY SIMILARITY.
FT	DISULFID	853	862	BY SIMILARITY.
FT	DISULFID	865	875	BY SIMILARITY.
FT	DISULFID	878	887	BY SIMILARITY.
FT	DISULFID	880	894	BY SIMILARITY.
FT	DISULFID	897	906	BY SIMILARITY.
FT	DISULFID	909	925	BY SIMILARITY.
FT	DISULFID	928	944	BY SIMILARITY.
FT	DISULFID	930	955	BY SIMILARITY.
FT	DISULFID	957	966	BY SIMILARITY.
FT	DISULFID	969	984	BY SIMILARITY.
FT	DISULFID	987	1001	BY SIMILARITY.
FT	DISULFID	989	1008	BY SIMILARITY.
FT	DISULFID	1011	1020	BY SIMILARITY.
FT	DISULFID	1023	1036	BY SIMILARITY.
FT	DISULFID	1098	1108	BY SIMILARITY.
FT	DISULFID	1117	1126	BY SIMILARITY.
FT	DISULFID	1129	1141	BY SIMILARITY.
FT	DISULFID	1144	1156	BY SIMILARITY.
FT	DISULFID	1146	1163	BY SIMILARITY.
FT	DISULFID	1165	1174	BY SIMILARITY.
FT	DISULFID	1171	1188	BY SIMILARITY.
FT	DISULFID	1191	1191	INTERCHAIN (PROBABLE).
FT	DISULFID	1194	1194	INTERCHAIN (PROBABLE).
FT	DISULFID	1198	1798	INTERCHAIN (PROBABLE).
FT	CARBOHYD	251	251	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	371	371	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1086	1086	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1250	1250	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1309	1309	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1349	1349	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	1500	1500	N-LINKED (GLCNAC. . .)
FT	SEQUENCE	1799 AA;	196352 MW;	1P28967A67AEDE33 CRC64;

Query Match	29.4%;	Score 60;	DB 1;	Length 1799;
Best Local Similarity	29.8%;	Pred.No. 7.6;		
Matches 14;	Conservative 5;	Mismatches 8;	Indels 20;	Gaps

QY	5	YQKQCHM-----	-----GTHCHSD-GPRGVIPPRC 32
DB	1111	FTGQCHCHAGFGGRTCSQCQLYWGDPGLQCRACDCDPRG-IDKPOC 1156	

RESULT 2					
ID	LMB2	RAT	STANDARD;	PRT;	1801 AA.
AC	P15800;				
DT	01-APR-1990	(Rel. 14, Created)			
DT	01-APR-1990	(Rel. 14, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Laminin beta-2 chain precursor (S-laminin) (Laminin chain B3).				
DS	LAMB2.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata;				
OC	Mammalia; Eutheria; Rodentia;				
OX	NCBI_TaxId=10116;				
RN	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RC	MEDLINE=89159410; PubMed=2922051;				
RA	Hunter D.D.; Shah V.; Merlie J.P.;				
RT	"A laminin-like adhesive protein concentrated in the synaptic cleft				
RT	of the neuromuscular junction."				
RL	Nature 338:229-234(1989).				
CC	-I- FUNCTION: Binding to cells via a high affinity receptor, laminin				
CC	is thought to mediate the attachment, migration and organization				
CC	of cells into tissues during embryonic development by interacting				
CC	with other extracellular matrix components.				
CC	SUBUNIT: Laminin is a complex glycoprotein, consisting of three				
CC	different polypeptide chains (alpha, beta, gamma), which are bound				
CC	to each other by disulfide bonds into a cross-shaped molecule				

DISULFID	473	487	BY SIMILARITY.
DISULFID	476	494	BY SIMILARITY.
DISULFID	495	505	BY SIMILARITY.
DISULFID	508	522	BY SIMILARITY.
DISULFID	786	798	BY SIMILARITY.
DISULFID	788	805	BY SIMILARITY.
DISULFID	807	816	BY SIMILARITY.
DISULFID	819	831	BY SIMILARITY.
DISULFID	834	846	BY SIMILARITY.
DISULFID	836	853	BY SIMILARITY.
DISULFID	855	864	BY SIMILARITY.
DISULFID	867	877	BY SIMILARITY.
DISULFID	880	889	BY SIMILARITY.
DISULFID	882	896	BY SIMILARITY.
DISULFID	899	908	BY SIMILARITY.
DISULFID	911	927	BY SIMILARITY.
DISULFID	930	946	BY SIMILARITY.
DISULFID	932	957	BY SIMILARITY.
DISULFID	959	968	BY SIMILARITY.
DISULFID	971	986	BY SIMILARITY.
DISULFID	989	1003	BY SIMILARITY.
DISULFID	991	1010	BY SIMILARITY.
DISULFID	1013	1022	BY SIMILARITY.
DISULFID	1025	1038	BY SIMILARITY.
DISULFID	1098	1110	BY SIMILARITY.
DISULFID	1100	1117	BY SIMILARITY.
DISULFID	1119	1128	BY SIMILARITY.
DISULFID	1131	1143	BY SIMILARITY.
DISULFID	1146	1158	BY SIMILARITY.
DISULFID	1148	1165	BY SIMILARITY.
DISULFID	1167	1176	BY SIMILARITY.
DISULFID	1179	1190	BY SIMILARITY.
DISULFID	1193	1193	INTERCHAIN (PROBABLE).
DISULFID	1196	1196	INTERCHAIN (PROBABLE).
DISULFID	1800	1800	INTERCHAIN (PROBABLE).
CARBOHYD	251	251	N-LINKED (GLCNAC. .)
CARBOHYD	371	371	N-LINKED (GLCNAC. .)
CARBOHYD	1088	1088	N-LINKED (GLCNAC. .)
CARBOHYD	1252	1252	N-LINKED (GLCNAC. .)
CARBOHYD	1311	1311	N-LINKED (GLCNAC. .)
CARBOHYD	1351	1351	N-LINKED (GLCNAC. .)
CARBOHYD	1502	1502	N-LINKED (GLCNAC. .)
SEQUENCE	1801 AA;	196473 MW;	97AEF32PF8F31FA75 CRC64;
Query Match	29.48;	Score 60;	DB 1; Length 1801;
Best Local Similarity	29.84;	Pred. No. 7.6;	
Matches 14;	Conservative 5;	Mismatches 8;	Indels 20;
QY	5 YQKCHCHM-----	-----GTHCHSSD-GPRGVIEPRC 32	
Db	1113 FTGQCHAGFGGTCSEQELHWGDFGLQCRACDDPRG-IDKPOC 1158		
RESULT 3			
YH1 EBV			
ID1 YH1 EBV			
AC	P03181.	STANDARD;	PRT; 660 AA.
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	15-JUL-1998 (Rel. 36, Last annotation update)		
DE	Hypothetical BHLF1 protein.		
OS	Epstein-barr virus (strain B95-8) (Human herpesvirus 4).		
OC	Viruses; daDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae; Lymphocryptovirus.		
OC	NCB1_Taxid=10377;		
ON	[1]		
RX	SEQUENCE FROM N.A.		
RP	MEDLINE=84270667; PubMed=6087149;		
RA	Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,		
RA	Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,		
RA	Tuffnell P.S., Barrall B.G.;		
RT	"DNA sequence and expression of the B95-8 Epstein-Barr virus gene		
RL	Nature 310:207-211 (1984).		

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01555; -; NOT_ANNOTATED_CDS.
CC DR PIR; A03742; Q0B3.
CC KW Hypothetical protein; Early protein; Repeat.
CC FT DOMAIN 149 648 4 X 125 AA TANDEN REPEATS.
CC FT REPEAT 149 273 1.
CC FT REPEAT 274 398 2.
CC FT REPEAT 399 523 3.
CC FT REPEAT 524 648 4.
CC SQ SEQUENCE 660 AA; 66244 MW; 86DA1D67A37152A2 CRC64;
CC -----
CC Query Match 29.2%; Score 59.5; DB 1; Length 660;
CC Best Local Similarity 57.1%; Pred. No. 3 4;
CC Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
CC -----
CC QY 13 MGTCHSSDPRGVPEPRCP 33
CC Db 1 MGTFCQARGPR-TTLPHP 20
CC -----
CC RESULT 4
CC NT02 HUMAN STANDARD; PRT; 2471 AA.
CC ID Q04721; Q99734; Q9H240;
CC AC Q04721; Q99734; Q9H240;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (hM2).
CC GN NOTCH2.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Brain;
CC RA Blaumüller C.M., Mann R.S.;
CC RT "Complete human notch 2 (hM2) cDNA sequence.";
CC RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC [2]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Breast tumor;
CC RA Correa R.G., Canargo A.A., Moreira E.S., Simpson A.J.G.;
CC RT "Human Notch2, a novel member of cell-fate determining NOTCH
CC family.";
CC RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC [3]
CC RP SEQUENCE OF 967-1229 FROM N.A.
CC RC TISSUE=T-cell;
CC RA Lenasson I., Devaux C., Meenard J.M.;
CC RT "Partial sequence of EGF-like repeat domain of human Notch2 mRNA.";
CC RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC [4]
CC RP SEQUENCE OF 1810-2447 FROM N.A.
CC RC TISSUE=Brain;
CC RX MEDLINE=93265135; PubMed=1303260;
CC RA Stifani S., Blaumüller C.M., Redhead N.J., Hill R.E.,
CC RA Artavanis-Tsakonas S.;
CC RT "Human homologs of a Drosophila enhancer of split gene product define
CC a novel family of nuclear proteins.";
CC RL Nat. Genet. 2:119-127(1992).
CC [5]
CC RP POST-TRANSLATIONAL PROCESSING.
CC RX MEDLINE=97386453; PubMed=9244302;
CC RA Blaumüller C.M., Qi H., Zagouras P., Artavanis-Tsakonas S.;

```

```

RT "Intracellular cleavage of Notch leads to a heterodimeric receptor on
RT the plasma membrane.";
RL Cell 90:281-291(1997).
RN [6]
RX IDENTIFICATION OF LIGANDS.
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor.";
RL Am. J. Pathol. 154:785-794(1999).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity).
CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -1- TISSUE SPECIFICITY: Expressed in the brain, heart, kidney, lung,
CC skeletal muscle and liver.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TW) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to the NOTCH family.
CC -1- SIMILARITY: Contains 35 EGF-like domains.
CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
CC -1- SIMILARITY: Contains 6 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF308601; AAA36377.2; -
CC EMBL; AF315356; AAG37073.1; -
CC EMBL; U77493; AAB19224.1; -
CC HSSP; P00740; 1EDM
CC Genew; HGNC:7882; NOTCH2.
CC MIM; 600275; -
CC DR InterPro; IPR002110; ANK.
CC DR InterPro; IPR000152; Asx_hydroxyl_S.
CC DR InterPro; IPR000742; EGF_2.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR001438; EGF_II.
CC DR InterPro; IPR006209; EGF_like.
CC DR InterPro; IPR002049; Laminin_EGF.
CC DR InterPro; IPR008297; Notch.
CC DR InterPro; IPR008000; Notch_dom.
CC DR Pfam; PF00023; ank; 6.
CC DR Pfam; PF00008; EGF; 35.
CC DR Pfam; PF00066; notch; 2.
CC DR PIRSF; PIRSF002279; Notch; 1.
CC DR PRINTS; PR00010; EGFLOCD.
CC DR PRINTS; PR00011; EGFAMININ.
CC DR PRINTS; PR01452; NOTCH.
CC DR SMART; SM00248; ANK; 6.
CC DR SMART; SM00179; EGF_CA; 23.

```

DR SMART; SMO0004; NL; 2.
 DR PROSITE; PS00297; ANK REP REGION; 1.
 DR PROSITE; PS00088; ANK REPEAT; 4.
 DR PROSITE; PS00010; ASX HYDROXYL; 22.
 DR PROSITE; PS00022; EGF 1; 34.
 DR PROSITE; PS01186; EGF 2; 29.
 DR PROSITE; PS00026; EGF 3; 35.
 DR PROSITE; PS01187; EGF CA; 22.
 KW Receptor; Transcription regulation; Activator; Differentiation;
 KW Developmental protein; Receptor; ANK repeat; EGF-like domain;
 KW Transmembrane; Glycoprotein; Signal; Phosphorylation.
 FT SIGNAL 1 25
 FT CHAIN 26 2471
 FT CHAIN 1666 2471
 FT CHAIN 1697 2471
 FT DOMAIN 26 1677
 FT TRANSMEM 1678 1698
 FT DOMAIN 1699 2471
 FT DOMAIN 26 63
 FT DOMAIN 64 102
 FT DOMAIN 105 143
 FT DOMAIN 144 180
 FT DOMAIN 182 219
 FT DOMAIN 221 258
 FT DOMAIN 260 296
 FT DOMAIN 298 336
 FT DOMAIN 338 374
 FT DOMAIN 375 413
 FT DOMAIN 415 454
 FT DOMAIN 456 492
 FT DOMAIN 494 530
 FT DOMAIN 532 568
 FT DOMAIN 570 605
 FT DOMAIN 607 643
 FT DOMAIN 645 680
 FT DOMAIN 682 718
 FT DOMAIN 720 755
 FT DOMAIN 757 793
 FT DOMAIN 795 831
 FT DOMAIN 833 871
 FT DOMAIN 873 909
 FT DOMAIN 911 947
 FT DOMAIN 949 985
 FT DOMAIN 987 1023
 FT DOMAIN 1025 1061
 FT DOMAIN 1063 1099
 FT DOMAIN 1101 1147
 FT DOMAIN 1149 1185
 FT DOMAIN 1187 1223
 FT DOMAIN 1225 1262
 FT DOMAIN 1264 1302
 FT DOMAIN 1304 1343
 FT DOMAIN 1374 1412
 FT REPEAT 1420 1456
 FT REPEAT 1503 1535
 FT REPEAT 1827 1871
 FT REPEAT 1875 1905
 FT REPEAT 1909 1939
 FT REPEAT 1943 1972
 FT REPEAT 1976 2005
 FT REPEAT 2009 2038
 FT DOMAIN 1645 1648
 FT DOMAIN 1994 1997
 FT DOMAIN 2426 2429
 FT DISULFID 28 41
 FT DISULFID 35 51
 FT DISULFID 53 62
 FT DISULFID 68 79
 FT DISULFID 73 90
 FT DISULFID 92 101
 FT DISULFID 109 121

FT DISULFID 115 131 BY SIMILARITY.
 FT DISULFID 133 142 BY SIMILARITY.
 FT DISULFID 148 159 BY SIMILARITY.
 FT DISULFID 153 168 BY SIMILARITY.
 FT DISULFID 170 179 BY SIMILARITY.
 FT DISULFID 186 198 BY SIMILARITY.
 FT DISULFID 192 207 BY SIMILARITY.
 FT DISULFID 209 218 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.
 FT DISULFID 230 246 BY SIMILARITY.
 FT DISULFID 248 257 BY SIMILARITY.
 Query Match 28.7%; Score 58.5; DB 1; Length 2471;
 Best Local Similarity 42.3%; Pred. No. 16;
 Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;
 Qy 7 GKCHCHMGTHC-HSSDGPGRGVPEPR 31
 Db 1347 GQVCKRGQCVHTASGPRCFPSPR 1372
 RESULT 5
 SMR1 MOUSE
 ID SMR1_MOUSE STANDARD; PRT; 147 AA.
 AC Q61900;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Submaxillary gland androgen regulated protein 1 precursor (Salivary protein MSGI).
 DE SMR1 OR MSGI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN-BALB/c; TISSUE-Submaxillary gland;
 RX MEDLINE=94252564; PubMed=8194749;
 RA Trautwein R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shennen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gurnatne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May play a role in protection or detoxification.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Secreted into saliva by submaxillary gland.
 CC -!- SIMILARITY: BELONGS TO THE PROL1/PROL3 FAMILY.

```

Query Match      26.2%  Score 53.5;  DB 1;  Length 2768;
Best local similarity 31.2%  Pred. No. 77;
Matches 15;  Conservative 2;  Mismatches 12;  Indels 19;  Gaps 3;

5 YQKGC-----HCHMGT-HCHSSDGRGVIP-----PRCP 33
      ||| ||| ||| ||| ||| ||| ||| ||| |||
1027 YMPQCDAGSWGPEVQCHAGTGCWCVDKGGFPGSLTARSLQIQCP 1074
      ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

FT DISULFID 378 411 BY SIMILARITY.
FT DISULFID 383 396 BY SIMILARITY.
FT DISULFID 398 403 BY SIMILARITY.
FT DISULFID 417 422 BY SIMILARITY.
FT DISULFID 419 450 BY SIMILARITY.
FT DISULFID 424 433 BY SIMILARITY.
FT DISULFID 435 442 BY SIMILARITY.
FT DISULFID 456 461 BY SIMILARITY.
FT DISULFID 458 504 BY SIMILARITY.
FT DISULFID 463 473 BY SIMILARITY.
FT DISULFID 476 479 BY SIMILARITY.
FT DISULFID 483 492 BY SIMILARITY.
FT DISULFID 489 561 BY SIMILARITY.
FT DISULFID 508 537 BY SIMILARITY.
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 577 577
SQ SEQUENCE 577 AA; 62298 MW; B83B468C83EDCCF9 CRC64;

Query Match 25.7%; Score 52.5; DB 1; Length 577;
Best Local Similarity 34.2%; Pred. No. 23;
Matches 13; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

QY 4 YQGKCHC---HMGTHCH-----SSDGRGVIPPRC 32
DB 333 YQGVACNPGHMGPHCEGCEGDTLSTDSKETPDHPC 370

RESULT 9
ZN44_HUMAN STANDARD; PRT; 589 AA.
AC P15621; P17018; Q9ULZ7;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 44 (Zinc finger protein K0X7) (Gonadotropin
DE inducible transcription repressor-2) (GIOT-2).
GN ZNF44 OR K0X7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Okada T.; Mizutani T.; Miyamoto K.;
RT "Identification and characterization of novel zinc finger proteins in
RT the human ovary."
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]_
RP SEQUENCE OF 355-518 FROM N.A.
RA Ammendola S.; Ciliberto G.;
RN Submitted (AUG-1989) to the EMBL/GenBank/DBJ databases.
RN [3]_
RP SEQUENCE OF 365-420 FROM N.A.
RC TISSUE=Lymphoid;
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RT cells."
RL New Biol. 2:363-374 (1990).
CC -1- FUNCTION: May be involved in transcriptional regulation.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB021642; BAA86988.1; -;
CC EMBL; X16281; CAA34357.1; -;
CC EMBL; X52338; CAA36564.1; -;
CC PIR; D56409; S06779.
CC HSP; P08048; Z2NF.
CC Genew; HGNC:13110; ZNF44.
CC MIM; 194542; -;
CC TRANSFAC; T04987; -;
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR007087; Znf_C2H2.
CC InterPro; IPR007086; Znf_C2H2_sub.
CC Pfam; PF01352; KRAB; 1.
CC Pfam; PF00096; zf-C2H2; 15.
CC PRINTS; PR00048; ZINC_FINGER.
CC ProDom; PD000003; Znf_C2H2; 7.
CC SMART; SM00349; KRAB; 1.
CC SMART; SM00355; Znf_C2H2; 15.
CC PROSITE; PS00805; KRAB; 1.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 16.
CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
CC Nuclear protein; Repeat.
CC DOMAIN 4 90 KRAB.
CC ZN_FING 141 163 C2H2-TYPE (ATYPICAL).
CC ZN_FING 197 219 C2H2-TYPE.
CC ZN_FING 225 247 C2H2-TYPE.
CC ZN_FING 253 275 C2H2-TYPE.
CC ZN_FING 281 303 C2H2-TYPE.
CC ZN_FING 309 331 C2H2-TYPE.
CC ZN_FING 337 359 C2H2-TYPE.
CC ZN_FING 365 387 C2H2-TYPE.
CC ZN_FING 393 415 C2H2-TYPE.
CC ZN_FING 421 443 C2H2-TYPE.
CC ZN_FING 476 498 C2H2-TYPE.
CC ZN_FING 504 526 C2H2-TYPE.
CC ZN_FING 532 554 C2H2-TYPE.
CC ZN_FING 560 582 C2H2-TYPE.
CC ZN_FING 510 518 GKAFSRFSY -> AKPSVSVSVT (IN REF. 2).
CC SEQUENCE 589 AA; 68102 MW; 24F3EA86617E7FDD CRC64;
CC
CC Query Match 25.7%; Score 52.5; DB 1; Length 589;
CC Best Local Similarity 45.5%; Pred. No. 23;
CC Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;
CC
CC QY 7 GKCHCHMGTH-----CHSSDGP 23
CC DB 315 GKAFCHLGSFQRHMHSGDGP 336
CC
CC RESULT 10
CC ITB5_HUMAN STANDARD; PRT; 799 AA.
CC ID ITB5_HUMAN
CC AC P18084;
CC DT 01-NOV-1990 (Rel. 16, Created)
CC DT 01-NOV-1990 (Rel. 16, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Integrin beta-5 precursor.
CC GN ITGB5.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.

RC TISSUE=Thymic epithelium;
RX MEDLINE=90228356; PubMed=2328726;
RA Ramaswamy H., Henler M.E.;
RT "Cloning, primary structure and properties of a novel human integrin
beta subunit";
RL EMBO J. 9:1561-1568(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90319111; PubMed=2371275;
RA Suzuki S., Huang Z.S., Tanihara H.;
RT "Cloning of an integrin beta subunit exhibiting high homology with
integrin beta 3 subunit";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5354-5358(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009141; PubMed=2211615;
RA McLean J.W., Vestal D.J., Cheres D.A., Bodary S.C.;
RT "cDNA sequence of the human integrin beta 5 subunit";
RL J. Biol. Chem. 265:17126-17131(1990).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Murty D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: INTEGRIN ALPHA-VBETA-5 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGAND.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5
CC -1- ASSOCIATES WITH ALPHA-V.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the integrin beta chain family.
CC -1- SIMILARITY: Contains 2 VNFA-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53002; CAA37188.1; --
CC EMBL; M35011; AAA52707.1; --
CC EMBL; J05633; AAA59183.1; --
CC EMBL; BC006541; AA06541.1; --
CC PIR; A38308; A38308.
CC HSP; P05106; IJUV2.
CC Genew; HGNC:6160; ITGB5.
CC MIM; 147561; --
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR002369; Integrin_B.
CC InterPro; IPR001169; Integrin_beta_C.
CC InterPro; IPR003659; Flexin-like.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF00362; integrin_B; 1.

DR PRINTS; PRO1186; INTEGRIN.
DR ProDom; PD001811; Integrin_B; 1.
DR SMART; SM00187; INB; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00243; INTEGRIN_BETA; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;
KW Repeat; Signal.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 799 INTEGRIN BETA-5.
FT DOMAIN 24 719 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 720 742 POTENTIAL.
FT DOMAIN 743 799 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 136 378 VNFA-LIKE.
FT DOMAIN 465 630
FT REPEAT 465 512 I
FT REPEAT 513 554 II.
FT REPEAT 555 593 III.
FT REPEAT 594 630 IV.
FT DISULFID 28 463 BY SIMILARITY.
FT DISULFID 36 46 BY SIMILARITY.
FT DISULFID 39 75 BY SIMILARITY.
FT DISULFID 49 64 BY SIMILARITY.
FT DISULFID 202 211 BY SIMILARITY.
FT DISULFID 259 300 BY SIMILARITY.
FT DISULFID 401 413 BY SIMILARITY.
FT DISULFID 433 682 BY SIMILARITY.
FT DISULFID 461 465 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
FT DISULFID 484 522 BY SIMILARITY.
FT DISULFID 489 498 BY SIMILARITY.
FT DISULFID 500 513 BY SIMILARITY.
FT DISULFID 528 533 BY SIMILARITY.
FT DISULFID 530 563 BY SIMILARITY.
FT DISULFID 535 548 BY SIMILARITY.
FT DISULFID 550 555 BY SIMILARITY.
FT DISULFID 569 574 BY SIMILARITY.
FT DISULFID 571 602 BY SIMILARITY.
FT DISULFID 576 585 BY SIMILARITY.
FT DISULFID 587 594 BY SIMILARITY.
FT DISULFID 608 613 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 615 625 BY SIMILARITY.
FT DISULFID 628 631 BY SIMILARITY.
FT DISULFID 635 644 BY SIMILARITY.
FT DISULFID 641 714 BY SIMILARITY.
FT DISULFID 661 690 BY SIMILARITY.
FT CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 552 552 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 654 654 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 705 705 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 193 193 T -> A (IN REF. 2).
FT CONFLICT 645 645 L -> P (IN REF. 3).
FT CONFLICT 790 792 MISSING (IN REF. 2).
SQ SEQUENCE 799 AA; 88053 MW; D7E4727CA310512B CRC64;
Query Match 25.7%; Score 52.5; DB 1; Length 799;
Best Local Similarity 31.4%; Pred. No. 31;
Matches 11; Conservative 4; Mismatches 17; Indels 3; Gaps 1;
QY 1 STVYQKCHC---HMGTCHSSDGRGVIPERC 32
DB 479 SGTVGLCECPGLGLTRCECDGENQSVVQNL 513
RESULT 11
DMN_HUMAN

KW	Plasma; Hemostasis; Repeat; Cell adhesion; Signal.
FT	SIGNAL . 22 BY SIMILARITY.
FT	PROPEP . 23 763 BY SIMILARITY.
FT	CHAIN . 764 2813 VON WILLEBRAND FACTOR.
FT	DOMAIN . 35 179 VFWD 1.
FT	DOMAIN . 388 541 VFWD 2.
FT	DOMAIN . 764 787 AMINO-TERMINAL.
FT	DOMAIN . 788 833 E1.
FT	DOMAIN . 826 863 CX.
FT	DOMAIN . 867 1013 VFWD 3.
FT	DOMAIN . 1277 1453 VWFA 1.
FT	DOMAIN . 1498 1665 VWFA 2.
FT	DOMAIN . 1691 1871 VWFA 3.
FT	DOMAIN . 1950 2102 VFWD 4.
FT	DOMAIN . 2216 2361 E2.
FT	DOMAIN . 2255 2328 VWFC 1.
FT	DOMAIN . 2429 2495 VWFC 2.
FT	DOMAIN . 2580 2645 VWFC 3.
FT	DOMAIN . 2744 2812 CICK.
FT	SITE . 531 533 CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE . 698 700 CELL ATTACHMENT SITE (POTENTIAL).
FT	SITE . 2507 2509 CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID . 767 808 BY SIMILARITY.
FT	DISULFID . 776 804 BY SIMILARITY.
FT	DISULFID . 810 821 BY SIMILARITY.
FT	DISULFID . 867 996 BY SIMILARITY.
FT	DISULFID . 869 1031 BY SIMILARITY.
FT	DISULFID . 898 993 BY SIMILARITY.
FT	DISULFID . 914 921 BY SIMILARITY.
FT	DISULFID . 1060 1084 BY SIMILARITY.
FT	DISULFID . 1071 1111 BY SIMILARITY.
FT	DISULFID . 1089 1091 BY SIMILARITY.
FT	DISULFID . 1153 1165 BY SIMILARITY.
FT	DISULFID . 1149 1169 BY SIMILARITY.
FT	DISULFID . 1126 1130 BY SIMILARITY.
FT	DISULFID . 1136 1199 BY SIMILARITY.
FT	DISULFID . 1234 1237 BY SIMILARITY.
FT	DISULFID . 1272 1458 BY SIMILARITY.
FT	DISULFID . 1669 1670 BY SIMILARITY.
FT	DISULFID . 1686 1872 BY SIMILARITY.
FT	DISULFID . 1879 1904 BY SIMILARITY.
FT	DISULFID . 1899 1940 OR 1942 (BY SIMILARITY).
FT	DISULFID . 1972 2123 BY SIMILARITY.
FT	DISULFID . 1950 2085 BY SIMILARITY.
FT	DISULFID . 1927 2088 BY SIMILARITY.
FT	DISULFID . 1993 2001 BY SIMILARITY.
FT	DISULFID . 2724 2774 BY SIMILARITY.
FT	DISULFID . 2739 2788 BY SIMILARITY.
FT	DISULFID . 2750 2804 BY SIMILARITY.
FT	DISULFID . 2754 2806 BY SIMILARITY.
FT	DISULFID ? 2811 BY SIMILARITY.
FT	CARBOHYD . 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 666 666 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 857 857 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 1231 1231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 1515 1515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 1574 1574 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2223 2223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2290 2290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2357 2357 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2400 2400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2546 2546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2585 2585 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD . 2790 2790 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT . 55 55 G -> E (IN REF. 2).
FT	CONFLICT . 70 70 V -> I (IN REF. 3).
FT	CONFLICT . 266 266 A -> G (IN REF. 2).
FT	CONFLICT . 280 280 I -> V (IN REF. 2).
FT	CONFLICT . 409 411 VCH -> ICQ (IN REF. 2).
FT	CONFLICT . 994 994 G -> A (IN REF. 1).
FT	CONFLICT . 1021 1021 F -> L (IN REF. 2).

FT CONFLICT 2381 2381 L -> P (IN REF. 2).
 FT CONFLICT 2406 2406 P -> L (IN REF. 2).
 SQ SEQUENCE 2813 AA; 309716 MW; 5DP93E1E5E72F80C CRC64;
 Query Match 25.7%; Score 52.5; DB 1; Length 2813;
 Best Local Similarity 28.6%; Pred. No. 1e+02;
 Matches 12; Conservative 5; Mismatches 10; Indels 15; Gaps 2;
 QY 4 YQGGK-----CHCHMG-THCHSSDPGRGVPEP 30
 710 YDGEIPQBEDIFSDHHTWCYCDGFMHCTTSGGLGSLPNP 751
 Db
 RESULT 13
 LMB2 HUMAN STANDARD; PRT; 1798 AA.
 ID LMB2 HUMAN STANDARD; PRT; 1798 AA.
 AC P55268; Q16321;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Laminin beta-2 chain precursor (S-laminin) (Laminin B1s chain).
 GN LAMB2 OR LAMS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95213013; PubMed=7698745;
 RA Wever U.M., Gerecke D.R., Durkin M.E., Kurtz K.S., Mattei M.-G.,
 RA Champliand M.F., Burgesson R.E., Albrechtsen R.;
 RT "Human beta 2 chain of laminin (formerly S chain): cDNA cloning,
 RT chromosomal localization, and expression in carcinomas.";
 RL Genomics 24:243-252(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95316263; PubMed=7795887;
 RA Iivanainen A., Vuolteenaho R., Sainio K., Eddy R., Shows T.B.,
 RA Sariola H., Tryggvason K.;
 RT "The human laminin beta 2 chain (S-laminin): structure, expression in
 RT fetal tissues and chromosomal assignment of the LAMB2 gene.";
 RL Matrix Biol. 14:489-497(1995).
 CC -!- FUNCTION: Binding to cells via a high affinity receptor, laminin
 CC is thought to mediate the attachment, migration and organization
 CC of cells into tissues during embryonic development by interacting
 CC with other extracellular matrix components.
 CC -!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
 CC different polypeptide chains (alpha, beta, gamma), which are bound
 CC to each other by disulfide bonds into a cross-shaped molecule
 CC comprising one long and three short arms with globules at each
 CC end. The beta-2 chain is a subunit of laminin-3 (S-laminin),
 CC laminin-4 (S-merosin), and laminin-7 (KS-laminin).
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: Found in the basement membranes (major
 CC component). S-laminin is concentrated in the synaptic
 CC cleft of the neuromuscular junction.
 CC -!- DOMAIN: The alpha-helical domains I and II are thought to interact
 CC with other laminin chains to form a coiled coil structure.
 CC -!- DOMAIN: Domains VI and IV are globular.
 CC -!- SIMILARITY: Contains 1 laminin N-terminal domain.
 CC -!- SIMILARITY: Contains 13 laminin EGF-like domains.
 CC -!- SIMILARITY: Contains 1 laminin IV domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z68155; CAA92279.1; --
 DR EMBL; Z68156; CAA92279.1; JOINED.

DR EMBL; X79683; CAA56130.1; --
 DR EMBL; S77512; AAB34682.2; --
 DR PIR; S53869; S53869.
 DR HSSP; P02468; IKLO.
 DR Genew; HGNC:8487; LAMB2.
 DR MIM; 150325; --
 DR GO; GO:0005605; C:basal lamina; TAS.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR002049; Laminin_EGF.
 DR InterPro; IPR008211; LamNT.
 DR Pfam; PF00053; laminin_EGF; 13.
 DR Pfam; PF00055; laminin_Nterm; 1.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR SMART; SMO0180; EGF_Lam; 13.
 DR SMART; SMO0136; LamNT; 1.
 DR PROSITE; PS00022; EGF_1; 10.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 12.
 KW Glycoprotein. Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 32
 FT CHAIN 33 1798
 FT LAMININ BETA-2 CHAIN.
 FT DOMAIN 33 280
 FT LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 283 346
 FT LAMININ EGF-LIKE 1.
 FT DOMAIN 347 409
 FT LAMININ EGF-LIKE 2.
 FT DOMAIN 410 469
 FT LAMININ EGF-LIKE 3.
 FT DOMAIN 470 521
 FT LAMININ EGF-LIKE 4.
 FT DOMAIN 522 552
 FT LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 553 781
 FT LAMININ DOMAIN IV.
 FT DOMAIN 783 830
 FT LAMININ EGF-LIKE 6.
 FT DOMAIN 831 876
 FT LAMININ EGF-LIKE 7.
 FT DOMAIN 877 926
 FT LAMININ EGF-LIKE 8.
 FT DOMAIN 927 985
 FT LAMININ EGF-LIKE 9.
 FT DOMAIN 986 1037
 FT LAMININ EGF-LIKE 10.
 FT DOMAIN 1038 1094
 FT LAMININ EGF-LIKE 11.
 FT DOMAIN 1095 1142
 FT LAMININ EGF-LIKE 12.
 FT DOMAIN 1143 1189
 FT LAMININ EGF-LIKE 13.
 FT DOMAIN 1190 1409
 FT DOMAIN II.
 FT DOMAIN 1410 1442
 FT DOMAIN ALPHA.
 FT DOMAIN 1443 1798
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1253 1319
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1472 1526
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1577 1790
 FT BY SIMILARITY.
 FT DISULFID 283 292
 FT BY SIMILARITY.
 FT DISULFID 312 321
 FT BY SIMILARITY.
 FT DISULFID 324 344
 FT BY SIMILARITY.
 FT DISULFID 347 356
 FT BY SIMILARITY.
 FT DISULFID 349 374
 FT BY SIMILARITY.
 FT DISULFID 377 386
 FT BY SIMILARITY.
 FT DISULFID 389 407
 FT BY SIMILARITY.
 FT DISULFID 410 423
 FT BY SIMILARITY.
 FT DISULFID 412 438
 FT BY SIMILARITY.
 FT DISULFID 440 449
 FT BY SIMILARITY.
 FT DISULFID 452 467
 FT BY SIMILARITY.
 FT DISULFID 470 484
 FT BY SIMILARITY.
 FT DISULFID 472 491
 FT BY SIMILARITY.
 FT DISULFID 493 502
 FT BY SIMILARITY.
 FT DISULFID 505 519
 FT BY SIMILARITY.
 FT DISULFID 783 795
 FT BY SIMILARITY.
 FT DISULFID 785 802
 FT BY SIMILARITY.
 FT DISULFID 804 813
 FT BY SIMILARITY.
 FT DISULFID 816 828
 FT BY SIMILARITY.
 FT DISULFID 831 843
 FT BY SIMILARITY.
 FT DISULFID 833 850
 FT BY SIMILARITY.
 FT DISULFID 852 861
 FT BY SIMILARITY.
 FT DISULFID 864 874
 FT BY SIMILARITY.
 FT DISULFID 877 886
 FT BY SIMILARITY.
 FT DISULFID 879 893
 FT BY SIMILARITY.
 FT DISULFID 896 905
 FT BY SIMILARITY.
 FT DISULFID 908 924
 FT BY SIMILARITY.
 FT DISULFID 927 943
 FT BY SIMILARITY.
 FT DISULFID 929 954
 FT BY SIMILARITY.

```
FT DISULFID 956 965 BY SIMILARITY.
FT DISULFID 968 983 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 988 1007 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1022 1035 BY SIMILARITY.
FT DISULFID 1095 1107 BY SIMILARITY.
FT DISULFID 1097 1114 BY SIMILARITY.
FT DISULFID 1116 1125 BY SIMILARITY.
FT DISULFID 1128 1140 BY SIMILARITY.
FT DISULFID 1143 1155 BY SIMILARITY.
FT DISULFID 1145 1162 BY SIMILARITY.
FT DISULFID 1164 1173 BY SIMILARITY.
FT DISULFID 1176 1187 BY SIMILARITY.
FT DISULFID 1190 1191 INTERCHAIN (PROBABLE).
FT DISULFID 1193 1193 INTERCHAIN (PROBABLE).
FT DISULFID 1797 1797 INTERCHAIN (PROBABLE).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 914 914 R -> G (IN REF. 2).
FT CONFLICT 1179 1179 G -> A (IN REF. 2).
SQ SEQUENCE 1798 AA; 196079 MW; 9555CF5B24850CB7 CRC64;

Query Match 25.5%; Score 52; DB 1; Length 1798;
Best Local Similarity 23.9%; Pred. No. 78;
Matches 11; Conservative 4; Mismatches 13; Indels 19; Gaps 1;

QY 5 YQKCHCHM-----GTHCHSSDPRGVPEPC 32
DB 1110 FTQCHCRAGFGGTCSECELHWGDPGLQCHACDCDSRGIDTPQC 1155

RESULT 14
HMRI RHIME
ID HMRI RHIME STANDARD; PRT; 147 AA.
AC P86378;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Heavy metal dependent transcription regulator 1.
GN HMRI1 OR HMR1 OR R20549 OR SNA1014
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Guzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
RC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR INVOLVED IN ACID TOLERANCE.
CC -!- BINDS COPPER (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- DOMAIN: It contains a N-terminal DNA binding region and a C-
CC terminal metal binding region (By similarity).
CC -!- SIMILARITY: Contains 1 HTH merR-type DNA-binding domain.
CC -!- CAUTION: We have changed a potential AGG for Arg-1 as ATG for Met-
CC 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE007245; AAK65207.1; ALT_SEQ.
CC InterPro; IPR000551; HTH_MerR.
CC Pfam; PF00376; merR; 1.
CC PRINTS; PR00040; HTHMERR.
CC SMART; SMO0422; HTH_MERR; 1.
CC PROSITE; PS00552; HTH_MERR_1; FALSE NEG.
CC PROSITE; PS00937; HTH_MERR_2; 1.
CC Transcription regulation; DNA-binding; Copper; Plasmid;
KW Complete proteome.
FT DOMAIN 1 70 HTH_MERR-TYPE.
FT DNA BIND 3 22 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 147 AA; 16415 MW; 3A447536A9C07A86 CRC64;

Query Match 25.2%; Score 51.5; DB 1; Length 147;
Best Local Similarity 36.4%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

QY 12 HMGTHCHSSDPRGVPEPC 33
DB 108 HLAHCHGDD-----RPEC 122

RESULT 15
WIFI BRARE
ID WIFI BRARE STANDARD; PRT; 378 AA.
AC Q9W6F9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE WNT inhibitory factor 1 precursor (WIF-1).
GN WIFI.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9215557; PubMed=10201374;
RA Hsieh J.-C., Kodjabachian L., Rebbert M.L., Rattner A.,
RA Smallwood P.M., Samos C.H., Nusse R., Dawid I.B., Nathans J.;
RT "A new secreted protein that binds to Wnt proteins and inhibits their
RT activities."
RL Nature 398:431-436(1999).
RC -!- FUNCTION: Binds to WNT proteins and inhibits their activities. May
CC be involved in mesoderm segmentation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Highly expressed in unsegmented paraxial
CC mesoderm.
CC -!- SIMILARITY: Contains 5 EGF-like domains.
CC -!- SIMILARITY: Contains 1 WIF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF122925; AAD25405.1; -.
CC PIR; B59180; B59180.
CC HSPF; P00740; 1EDM.
CC ZFIN; ZDB-GENE-990712-17; wifi.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; IEGF.
```

DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR003306; WIF.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF02019; WIF; 1.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00181; EGF; 5.
DR SMART; SM00469; WIF; 1.
DR PROSITE; PS00022; EGF_1; 5.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS00026; EGF_3; 4.
DR PROSITE; PS00914; WIF; 1.
KW Repeat; EGF-like domain; Signal; Developmental protein;
KW Wnt signaling pathway.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 378 WNT INHIBITORY FACTOR 1.
FT DOMAIN 36 174 WIF.
FT DOMAIN 175 206 EGF-LIKE 1.
FT DOMAIN 207 238 EGF-LIKE 2.
FT DOMAIN 239 270 EGF-LIKE 3.
FT DOMAIN 271 302 EGF-LIKE 4.
FT DOMAIN 303 334 EGF-LIKE 5.
FT DISULFID 175 184 POTENTIAL.
FT DISULFID 180 190 POTENTIAL.
FT DISULFID 196 198 POTENTIAL.
FT DISULFID 207 216 POTENTIAL.
FT DISULFID 212 222 POTENTIAL.
FT DISULFID 228 230 POTENTIAL.
FT DISULFID 239 248 POTENTIAL.
FT DISULFID 244 254 POTENTIAL.
FT DISULFID 260 262 POTENTIAL.
FT DISULFID 271 280 POTENTIAL.
FT DISULFID 276 286 POTENTIAL.
FT DISULFID 292 294 POTENTIAL.
FT DISULFID 303 312 POTENTIAL.
FT DISULFID 308 318 POTENTIAL.
FT DISULFID 324 326 POTENTIAL.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 378 AA; 41312 MW; 42FE9F70D948D1D8 CRC64;

Query Match 25.2%; Score 51.5; DB 1; Length 378;
Best Local Similarity 36.1%; Pred. No. 20;
Matches 13; Conservative 1; Mismatches 13; Indels 9; Gaps 2;

QY 6 QGKCHCHMGTH-----CHSSDGRGVIPER-RC 32
Db 289 RNKCKCKGYHGLCKSKAVCEPSCGAHGTCTVEPNRC 324

Search completed: February 26, 2004, 12:10:52
Job time : 5.80106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 14.0928 Seconds
(without alignments)
738.822 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVIPERCP 33

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	165	4	O95669
2	204	100.0	190	4	O14932
3	142	69.6	180	6	Q8MJ02
4	119	58.3	151	6	Q8MJ01
5	119	58.3	176	6	Q95JB8
6	67	32.8	192	11	Q8CG11
7	67	32.8	192	11	Q8W0M8
8	67	32.8	192	11	Q8CFD9
9	63.5	31.1	665	16	Q7WQL1
10	63.5	31.1	665	16	Q7WIN6
11	63.5	31.1	665	16	Q7VZW9
12	61	29.9	485	5	O97271
13	60	29.4	176	4	O95667
14	60	29.4	201	4	O14931
15	60	29.4	1799	11	Q8R0Y0
16	59	28.9	544	16	Q9X8Z2

17	58.5	28.7	116	4	Q9NT74
18	58	28.4	180	10	O82297
19	58	28.4	845	11	O61072
20	58	28.4	853	11	O80U94
21	57.5	28.2	198	12	O69020
22	57.5	28.2	881	5	Q9W0A0
23	56.5	27.7	921	11	O80T91
24	56.5	27.7	969	4	Q96KG6
25	55.5	27.2	1140	4	Q96KG7
26	55.5	27.2	1140	5	O8MLU9
27	55.5	27.2	2931	5	Q9W2C6
28	55	27.0	1664	5	Q9TVQ2
29	54.5	26.7	144	16	O8DU62
30	54.5	26.7	3078	5	Q26031
31	54	26.5	258	16	O89LM8
32	54	26.5	491	4	O8TEK2
33	54	26.5	546	11	O8OV70
34	54	26.5	747	11	O8VHF4
35	54	26.5	1004	11	O8CGA7
36	54	26.5	1034	11	O8VHL7
37	54	26.5	1034	11	O8VIK5
38	53.5	26.2	289	16	O87LT5
39	53	26.0	142	10	O22479
40	53	26.0	153	11	O8VHC9
41	53	26.0	297	12	O91BJ2
42	53	26.0	334	10	O9FTM4
43	53	26.0	473	2	Q8RJF1
44	53	26.0	597	5	Q19286
45	53	26.0	818	11	Q9Z0M6

ALIGNMENTS

RESULT 1

O95669 PRELIMINARY; PRT; 165 AA.

AC O95669;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE 1C7f.
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bouguieret L., Prieur S., Caterina D., Primas G., Perrot V., Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFKappaB family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J., Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H., Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [4]
 RP SEQUENCE FROM N.A.

Q9nt74 homo sapien
 O82297 arabidopsis
 O61072 mus musculus
 O80U94 mus musculus
 O69020 human herpe
 Q9W0A0 drosophila
 O80T91 mus musculus
 Q96KG6 homo sapien
 Q96KG7 homo sapien
 O8MLU9 drosophila
 Q9W2C6 drosophila
 Q9TVQ2 caenorhabdi
 Q8DJ62 synchococc
 Q26031 plasmodium
 O89LM8 bradyrhizob
 Q8TEK2 homo sapien
 O8OV70 mus musculus
 O8VHF4 mus musculus
 O8CGA7 mus musculus
 O8VHL7 mus musculus
 O8VIK5 mus musculus
 O87LT5 vibrio para
 O22479 santalum al
 O8VHC9 mus musculus
 Q91BJ2 spodoptera
 Q9FTM4 oryza sativ
 Q8RJF1 hyphomicrob
 Q19286 caenorhabdi
 Q9Z0M6 mus musculus

DR EMBL; Y14768; CAA75068.1; --
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;
Query Match 100.0%; Score 204; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. NO. 6.1e-21;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STVYQKCHCHMGTHCHSSDGRGVIPERCP 33
DB 133 STVYQKCHCHMGTHCHSSDGRGVIPERCP 165
RESULT 2
O14932
ID O14932 PRELIMINARY; PRT; 190 AA.
AC O14932;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE IC7 precursor (Natural killer cell receptor).
GN IC7 OR NKFP30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Spleen;
RX MEDLINE=96422187; PubMed=8824804;
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";
RL Genomics 31:215-222(1996).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=Spleen;
RX MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
RT Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392(1996).
RN [5]
SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RT TNF region.";
RL Immunogenetics 42:315-322(1995).
RN [6]
SEQUENCE FROM N.A.
RP TISSUE=Spleen;
RX MEDLINE=93208881; PubMed=7916655;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,

RL MEDLINE=93208881; PubMed=7916655;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chowing J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chowing J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
RT "Lymphotxin-beta: A new member of the TNF family that forms a
RT heteromeric complex with lymphotoxin on the cell surface.";
RL Cell 72:847-856(1993).
RN [5]
SEQUENCE FROM N.A.
RX MEDLINE=86016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotxin and tumor necrosis factor genes: structure,
RT homology and chromosomal localization.";
RL Nucleic Acids Res. 13:6361-6373(1985).
RN [6]
SEQUENCE FROM N.A.
RX MEDLINE=91086846; PubMed=1670638;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Page G.R.,
RA Riethmuller G., Weiss E.H.;
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
RT NcoI Polymorphism in the First Intron of the Human TNF-beta Gene
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
RT Level of TNF-beta Production.";
RL J. Exp. Med. 173:209-219(1991).
RN [7]
SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671667;
RA Abraham J.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RT "Haplotypic polymorphisms of the TNFB gene.";
RL Immunogenetics 33:50-53(1991).
RN [8]
SEQUENCE FROM N.A.
RX MEDLINE=94362679; PubMed=8081366;
RA Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
RT histocompatibility complex that encodes a potential new member of the
RT I kappa B family of proteins.";
RL Hum. Mol. Genet. 3:793-799(1994).
RN [9]
SEQUENCE FROM N.A.
RX MEDLINE=95324911; PubMed=7601445;
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
RT Nuclear RNA Helicase of the D-E-A-D Family.";
RL Genomics 26:210-218(1995).
RN [10]
SEQUENCE FROM N.A.
RX MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RT "Alternative splicing of the LST-1 gene located in the major
RT histocompatibility complex on human chromosome 6.";
RL DNA Seq. 8:155-160(1997).
RN [11]
SEQUENCE FROM N.A.
RX MEDLINE=98035983; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600(1997).
RN [12]
SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;
RA Shilina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
RT IRL and MICA genes at the centromeric end of the HLA class I
RL Genomics 47:372-382(1998).

RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced
 RT level of TNF-beta production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Stronger J., Spies T.;
 RT "The BART gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family.";
 RL Genomics 26:210-218(1995).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shihina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono K., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IkbL and MICA genes at the centromeric end of the HLA Class I
 RT region.";
 RL Genomics 47:372-382(1998).
 RN [15]
 RP SEQUENCE FROM N.A.

RC TISSUE=Lymphoid;
 RA Biassoni R., Pessino A., Malaespina A.;
 RT "NK-A1 activating NK receptor.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peripheral blood;
 RA Sato M., Yabe T., Ohashi J., Tsuchiya N., Hanaoka K., Tokunaga K.,
 RA Fuji T.;
 RT "Identification of two novel single nucleotide polymorphisms in the
 RT NKp30 gene in human natural killer cells.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031138; AAS86580.1; -
 DR EMBL; Y14768; CAA75065.1; -
 DR EMBL; AJ223153; CAB54004.1; -
 DR EMBL; AB055881; BAS78472.1; -
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW SIGNAL; Receptor.
 KW SIGNAL; 61 POTENTIAL.
 FT CHAIN 62 190 1C7
 SQ SEQUENCE 190 AA; 20640 MW; 1FAC919E20A6B18A CRC64;
 Query Match 100.0%; Score 204; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 7e-21;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33
 Db 158 STVYQKCHCHMGTHCHSSDGRGVIPRCP 190
 RESULT 3
 Q8MJ02 PRELIMINARY; PRT; 180 AA.
 AC Q8MJ02;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NKp30.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LaBonte M.L., Miller J., Letvin N.L.;
 RT "Molecular cloning of rhesus monkey NKp46 and NKp30 and identification
 RT of NKp46SD and NKp30S.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY035214; AAK63116.1; -
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;
 Query Match 69.6%; Score 142; DB 6; Length 180;
 Best Local Similarity 95.7%; Pred. No. 3.6e-12;
 Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 STVYQKCHCHMGTHCHSSDGP 23
 Db 158 STVYQKCHCHMGTHCHSSDGP 180

```
RESULT 4
Q8MJ01 ID Q8MJ01 PRELIMINARY; PRT; 151 AA.
AC Q8MJ01
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30v1ED.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and identification
of NKP46SD and NKP30S.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035216; AAK63118.1; -
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;

Query Match 58.3%; Score 119; DB 6; Length 151;
Best Local Similarity 94.7%; Pred. No. 5.3e-09;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHS 19
Db 133 STLYYQKCHCHMGTHCHS 151

RESULT 5
Q95JB8 ID Q95JB8 PRELIMINARY; PRT; 176 AA.
AC Q95JB8
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30 (NKP30v1).
GN NCR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541, 9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Rizzi M., Biassoni R.;
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in
Macaca fascicularis lymphoid cells.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification
of NKP46SD and NKP30S.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278389; CAC41081.1; -
DR EMBL; AY035215; AAK63117.1; -
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;

Query Match 32.8%; Score 67; DB 11; Length 192;
Best Local Similarity 71.4%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCHCHMG 179

RESULT 7
Q80WM8 ID Q80WM8 PRELIMINARY; PRT; 192 AA.
AC Q80WM8
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30.
GN NKP30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsieh C.L., Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RT "NK cells and transplantation.";
RL Transpl. Immunol. 9:111-114(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=LEW;
RC
```

```

RA Hsieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RA Martinez O.M., Krams S.M.;
RT "Identification, Cloning, and Characterization of a Novel Rat NK
RT Receptor, rNKP30: a Molecule Expressed in Liver Allografts.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL: AY273824; AAP13457.1; -.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG-like.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

Query Match
Best Local Similarity 32.8%; Score 67; DB 11; Length 192;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCLCHVG 179

RESULT 8
Q8CFD9 PRELIMINARY; PRT; 192 AA.
AC Q8CFD9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Icy protein precursor.
GN Icy.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=101116;
RN SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=NK cell;
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor Icy7.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL: A430418; CAD23066.1; -.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG-like.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 192 Icy7 PROTEIN.
SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;

Query Match
Best Local Similarity 32.8%; Score 67; DB 11; Length 192;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCLCHVG 179

RESULT 9
Q7WQL1 PRELIMINARY; PRT; 665 AA.
AC Q7WQL1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydatoin utilization protein B.
GN HYB OR BB0319.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.

```

```

OX NCBI_TaxID=518;
RN SEQUENCE FROM N.A.
RC STRAIN=RS50 / ATCC BAA-588;
MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Harris B., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Basham D., Bason N., Cherevach I.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR ENBL: BX640437; CAE30817.1; -.
KW Complete proteome.
SQ SEQUENCE 665 AA; 72398 MW; 5B244ED30DDA2762 CRC64;

```

```

Query Match
Best Local Similarity 31.1%; Score 63.5; DB 16; Length 665;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYQKCHCHMGTHCHSSDGRGVIP 28
Db 108 VYQGMVCFVGGHINTD-MGGAVP 132

```

```

RESULT 10
Q7WING PRELIMINARY; PRT; 665 AA.
AC Q7WING;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydatoin utilization protein B.
GN HYB OR BB0316.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Harris B., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Basham D., Bason N., Cherevach I.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR ENBL: BX640423; CAE40057.1; -.
KW Complete proteome.
SQ SEQUENCE 665 AA; 72497 MW; 5B2BEFC60DDE39D2 CRC64;

```

```

Query Match
Best Local Similarity 31.1%; Score 63.5; DB 16; Length 665;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYQKCHCHMGTHCHSSDGRGVIP 28
Db 108 VYQGMVCFVGGHINTD-MGGAVP 132

```

```

RESULT 11
Q7VZW9 ID Q7VZW9 PRELIMINARY; PRT; 665 AA.
AC Q7VZW9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypoxanthine utilization protein B.
GN HYU8 OR BP0754.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640413; CAE41060.1; -.
KW Complete proteome.
SQ SEQUENCE 665 AA; 72395 MW; F27204349A190508 CRC64;

Query Match 31.1%; Score 63.5; DB 16; Length 665;
Best Local Similarity 46.2%; Pred. NO. 1.5;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYQGGKCHCHMGTHCHSSDGRVIP 28
Db 108 VYQGGKCHCHMGTHCHSSDGRVIP 132

RESULT 12
O97271 ID O97271 PRELIMINARY; PRT; 485 AA.
AC O97271;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN MAL3P5.20.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gantley S., Gwilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jasek B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandram M.-A.,
RA Rutter S., Skelton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Barrell B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
RT falciparum.";
RL Nature 400:532-538 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368687;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,

```

```

RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C., Corton C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Doggett J.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Latke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandram M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL034556; CAB38982.3; -.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Hypothetical protein.
SQ SEQUENCE 485 AA; 57072 MW; 59C4467FC026C472 CRC64;

Query Match 29.9%; Score 61; DB 5; Length 485;
Best Local Similarity 25.0%; Pred. NO. 2.5;
Matches 11; Conservative 10; Mismatches 7; Indels 16; Gaps 2;

QY 4 YVQGGKCHCHMGTHCHSSDGRVIP 31
Db 156 HHKGRCSCKYVWTRGCHLGRFCRCHDQSHVPEGTQVVPFQ 199

RESULT 13
O95667 ID O95667 PRELIMINARY; PRT; 176 AA.
AC O95667;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7e.
GN 1C7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris P., Bouguetel L., Prieur S., Caterina D., Primas G., Perrot V.,
RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the NFKappaB
RT family within a 90 kilobase HLA class III segment.";
RL Nat. Genet. 3:137-145(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcoci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RT TNF region.";
RL Immunogenetics 42:315-322(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93208881; PubMed=7916655;

```

```

DR GO: GO:0003793, F:defense/immunity protein activity; NAS.
DR GO: GO:0006955, P:immune response; NAS.
DR GO: GO:0006954, P:inflammatory response; NAS.
DR InterPro: IPR003599, Ig.
DR InterPro: IPR007110, Ig-like.
DR Pfam: PF00047, Ig: 1.
DR SMART: SM00409, Ig: 1.
DR PROSITE: PS00835, IG_LIKE; 1.
SQ SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;

Query Match      23.4%; Score 60; DB 4; Length 176;
Best Local Similarity 43.2%; Pred. No. 1.3;
Matches 16; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

Qy 1 STVYQKCHCHMGTHCHSSDGR-----GVIEPRCP 33
Db 133 STVYQK-----CLTWKGRQLPAVVPAPLPP 161

RESULT 14
ID 014931 PRELIMINARY; PRT; 201 AA.
AC 014931;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DD 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7 precursor (NCR3 protein).
GN C17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Splicein;
RX MEDLINE=9422187; PubMed=8844804;
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."
RL Genomics 31:215-222(1996).
RN [2]
RS SEQUENCE FROM N.A.
RC TISSUE=Splicein;
RX MEDLINE=9422187; PubMed=8844804;
RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
RT Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RS SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA Lasky S., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
RT region.;"
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]
RS SEQUENCE FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris F., Souquerelet L., Prieur S., Caterina D., Primas G., Perrot
RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
RT "Dense Alu clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment.;"
RL Nat. Genet. 3:137-145(1993).
RN [5]
RS SEQUENCE FROM N.A.
RX MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allergic inflammatory factory-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.;"
RL Transplantation 61:1387-1392(1996).
RN [6]
RS SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;

```

RT "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208891; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.,
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a
 heteromeric complex with lymphotoxin on the cell surface.";
 Cell 72:847-856(1993).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.,
 RT "Human lymphotxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.,
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced
 RT level of TNF-beta production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.,
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:150-53(1991).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.,
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geoffroy C., Vaiman M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Stroninger J., Spies T.,
 RT "The B2M gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family.";
 RL Genomics 26:210-218(1995).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.,
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.,
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,

RA Kimura M., Inoko H.,
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IKS1 and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 RN [16]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [17]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Blood;
 RA Strausberg R.,
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031137; AAB86579.1; -;
 DR EMBL; AF129756; AAD18088.1; -;
 DR EMBL; Y14768; CAA75063.1; -;
 DR EMBL; BC025882; AAH52582.1; -;
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 KW Signal.
 FT CHAIN 1 61 POTENTIAL.
 FT SIGNAL 62 201 1C7.
 SQ SEQUENCE 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;
 Query Match 29.4%; Score 60; DB 4; Length 201;
 Best Local Similarity 43.2%; Pred. No. 1.5;
 Matches 16; Conservative 2; Mismatches 7; Indels 12; Gaps 2;
 Qy 1 STVYQKCHCHGTHCHSSDGP-----GVPEPRCP 33
 ||||| | : ||| : | : |
 Db 158 STVYQK-----CLTWKGPRLQPAVPEAPLP 186
 RESULT 15
 Q8R0Y0 PRELIMINARY; PRT; 1799 AA.
 ID Q8R0Y0 AC Q8R0Y0
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Laminin, beta 2.
 GN LAMB2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR ENBL; BC026051; AAH26051.1; -
DR MGD; MGI:99916; Lamb2.
DR GO; GO:0005578; C:extracellular matrix; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004623; F:phospholipase A2 activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0016042; P:lipid catabolism; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR02049; Laminin_EGF.
DR InterPro; IPR008211; LamNT.
DR InterPro; IPR01211; PhospholipaseA2.
DR Pfam; PF00053; laminin_EGF; 13.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00180; EGF_Lam; 13.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00022; EGF_1; 10.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 12.
DR PROSITE; PS00119; PA2_ASP; 1.
KW Laminin EGF-like domain.
SQ SEQUENCE 1799 AA; 196578 MW; 37CA24B9CDA0791F CRC64;

Query Match 29.4%; Score 60; DB 11; Length 1799;
Best Local Similarity 29.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

OY 5 YGKCHCHM-----GTHCHSSD-GPGVPEPRC 32
Db 1111 FTGQCHAGFGGRTSCQELYWGDPLQCRACDCDPRG-IDKPOC 1156

Search completed: February 26, 2004, 12:13:48
Job time : 17.0928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 20.4828 Seconds
(without alignments)
455.215 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_25Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003s.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	33	AAE02773	Human Nkp
2	204	100.0	190	AAE06401	Human B-c
3	204	100.0	190	AAE02769	Human Nkp
4	61.5	30.1	82	AAE035071	Arabidops
5	61.5	30.1	82	AAE11882	Arabidops
6	61.5	30.1	636	AAE66269	Rat TANGO
7	60	29.4	66	AAE010665	Human pol
8	60	29.4	201	AAE06403	Human B-c
9	60	29.4	1799	AAE03559	Mouse lam
10	60	29.4	1801	AAE05895	Rat lamin
11	60	29.4	1801	AAE60383	Rat Prote
12	59	28.9	57	AAE43863	Propionib
13	59	28.9	57	AAE40382	Propionib
14	59	28.9	2813	AAE54347	Canine vo
15	59	28.9	2813	AAE70557	Canine vo
16	58.5	28.7	77	AAE97268	Novel hum
17	58.5	28.7	148	AAE74028	Human pro
18	58.5	28.7	148	AAE74543	Human col
19	58.5	28.7	302	AAE03692	Novel hum
20	58.5	28.7	2469	AAE18207	Human MOL
21	58.5	28.7	2469	AAE18192	Human mol
22	58.5	28.7	2471	AAE027065	Human Not
23	58.5	28.7	2471	AAE06816	Human Not
24	58.5	28.7	2471	AAE79774	Human Not
25	58.5	28.7	2471	ABP72572	Human Not

ALIGNMENTS

RESULT 1

AAE02773
ID AAE02773 standard; peptide; 33 AA.

XX AAE02773;

AC AC (first entry)

DT 06-AUG-2001

XX Human Nkp30 receptor intracellular region sequence.

DE Human Nkp30 receptor intracellular region sequence.

XX Human, Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;

XX melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

XX immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

XX therapy; intracellular region.

XX Homo sapiens.

OS Homo sapiens.

XX WO200136630-A2.

PN 25-MAY-2001.

XX 15-NOV-2000; 2000WO-EP011697.

XX 15-NOV-1999; 99CA-02288307.

PR 15-NOV-1999; 99US-00440514.

XX (INNA-) INNATE PHARMA SAS.

XX (UYGE-) UNIV GENOVA.

PI Moretta A, Bottino C, Biassoni R;

XX WPI; 2001-329221/34.

XX Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.

XX Claim 1; Fig 7B; 83pp; English.

XX The invention relates to human Nkp30 receptor and its corresponding cDNA

XX molecule which is involved in natural cytotoxicity mediated by natural

XX killer (NK) cells and antibodies that identify the same. Nkp30 receptor

XX is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively

XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are

XX useful for detecting and/or quantifying the presence of NK cells in a

XX biological sample. The invention also provides kits for detecting and/or

XX quantifying the presence of NK cells, for the selective removal of NK

XX cells from a biological sample, for the positive and selective

XX purification of NK cells from a biological sample and for the in vitro

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GvT), and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. NKp30 antibodies are useful for identifying NKp30 natural ligands and allow assessment of the level of surface NKp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence NKp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the intracellular region of human NKp30 receptor

Sequence 33 AA;

Query Match 100.0%; Score 204; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 9.6e-18;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33
DB 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33

RESULT 2
AAV06401
ID AAV06401 standard; protein; 190 AA.
AC AAV06401;
XX
XX
XX 20-SBP-1999 (first entry)
XX
XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX
XX BMOG: B-cell myelin oligodendrocyte glycoprotein; human;
KW signal transduction; immunomodulator; antiinflammatory;
KW autoimmune disease; inflammation; gene therapy; diagnosis.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH 1..12 "leader peptide"
FT Peptide /note= "leader peptide"
FT Protein 13..190
FT Modified-site 42 /note= "mature protein"
FT Modified-site 68 /note= "N-glycosylated"
FT Modified-site 121 /note= "N-glycosylated"
FT Modified-site 139..162 /note= "N-glycosylated"
FT Domain /note= "transmembrane domain"
FT Peptide 166..190
FT /note= "alternatively spliced C-terminal end"
XX
XX WO9923867-A2.
XX
XX 20-MAY-1999.
XX
XX 05-NOV-1998; 98WO-US023826.
XX
XX 07-NOV-1997; 97US-0064761P.
XX
XX (BIOJ) BIOGEN INC.
XX
XX Browning J;
XX
XX WPI; 1999-418423/35.
DR N-PSDB; AAX59347.
XX
XX Novel B-cell myelin oligodendrocyte glycoproteins.

PS Claim 2; Page 42; 43pp; English.
XX
XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

XX Sequence 190 AA;

Query Match 100.0%; Score 204; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 5.4e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33
DB 158 STVYQKCHCHMGTHCHSSDGRGVIPRCP 190

RESULT 3
AAE02769
ID AAE02769 standard; protein; 190 AA.
XX
XX AAE02769;
XX
XX 06-AUG-2001 (first entry)
DT Human NKp30 receptor.
DE
XX
XX Human: NKp30 receptor; natural killer cell; cytostatic; antimicrobial;
KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
XX therapy.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 1..18
FT Peptide /label= "Signal peptide"
FT Protein 19..190
FT Region /label= "Mature_NKp30_receptor_protein"
FT /label= "Extracellular region"
FT /note= "Forms an immunoglobulin (Ig) V-like domain"
FT Modified-site 42 /note= "N-glycosylation site"
FT Modified-site 121 /note= "N-glycosylation site"
FT Region 139..157 /label= "N-glycosylation site"
FT Region 158..190 /label= "Transmembrane region"
FT Region /label= "Intracellular region"
XX
XX WO200136630-A2.
XX
XX 25-MAY-2001.
XX
XX 15-NOV-2000; 2000WO-EP011697.
XX
XX 15-NOV-1999; 99CA-02288307.
PR 15-NOV-1999; 99US-00440514.
XX
XX (INNA-) INNATE PHARMA SAS.
PA (UYGE-) UNIV GENOVA.
PA


```

PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 22-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 30.1%; Score 61.5; DB 3; Length 82;
Best Local Similarity 52.4%; Pred. No. 5.2;
Matches 11; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 3 VYQGRCHCHMGTHCHSSDGP 23
   : ||||| : |||||
Db 48 LPDGRKCH-HLKLHGSSHP 67

RESULT 6
AAB66269
ID AAB66269 standard; protein; 636 AA.
AC AAB66269;
XX
XX
DT 05-APR-2001 (first entry)
DE Rat TANGO 272 SEQ ID NO: 20.
XX
XX Membrane associated protein; secreted protein; human; mouse; rat;
KW INTERCEPT 340; MANGO 003; MANGO 347; TANGO 272; TANGO 295; TANGO 354;
KW TANGO 378; skeletal disorder; cardiovascular disorder; renal disorder;
KW haematopoietic disorder; neural disorder; hepatic disorder;
KW neoplastic disease.
XX
OS Rattus sp.
XX
XX WO200100673-A1.
XX
XX 04-JAN-2001.
XX
XX 29-JUN-2000; 2000WO-US018198.
XX
XX 30-JUN-1999; 99US-00345464.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Barnes TM, Fraser CC, Wrighton N, Myers P, Busfield SJ, Sharp JD;
XX WPI; 2001-050128/06.
XX
XX N-PSDB; AAF27791.
XX
XX Isolated secreted or transmembrane proteins are used for diagnosis and
XX treatment of neoplastic and hematopoietic disorders e.g. T cell
XX disorders, cancer and tumors.
XX
XX Claim 9; Page 238-240; 294pp; English.
XX
XX The present invention provides the protein and coding sequences for a
XX number of membrane associated and secreted proteins from human, mouse and
XX rat. The proteins are designated INTERCEPT 340, MANGO 003, MANGO 347,

CC TANGO 272, TANGO 295, TANGO 254 and TANGO 378. The proteins are all
CC involved in signal transduction and the sequences can be used in the
CC treatment of cardiovascular, renal, hepatic, neural, neoplastic, skeletal
CC and haematopoietic disorders
XX
SQ Sequence 636 AA;

Query Match 30.1%; Score 61.5; DB 4; Length 636;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSSDGPGRVIP---EPRC 32
   ||||| : |||||
Db 411 CQCHGATCPQDSCVCIFGTGPNC 437

RESULT 7
AAO10665
ID AAO10665 standard; protein; 66 AA.
XX
XX AAO10665;
AC AAO10665;
XX
XX 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 24557.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI90596.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 24557; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: the sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 66 AA;

Query Match 29.4%; Score 60; DB 4; Length 66;
Best Local Similarity 37.0%; Pred. No. 6.3;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

```

QY 4 YYQKCHCHMGTHCHSSDGRGVIP 30
 DB 27 HHHHCHCHYHSHHNSQHPPPPPP 53

RESULT 8
 ID AAY06403 standard; protein; 201 AA.
 AC AAY06403;
 XX
 XX 20-SEP-1999 (first entry)
 XX
 DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
 XX
 KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
 KW signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..12
 FT /note= "leader peptide"
 FT Protein 13..201
 FT /note= "mature protein"
 FT Modified-site 42
 FT /note= "N-glycosylated"
 FT Modified-site 68
 FT /note= "N-glycosylated"
 FT Modified-site 121
 FT /note= "N-glycosylated"
 FT Domain 139..162
 FT /note= "transmembrane domain"
 FT Peptide 166..201
 FT /note= "alternatively spliced C-terminal end"
 XX
 PN WO9923867-A2.
 XX
 XX 20-MAY-1999.
 XX
 XX 05-NOV-1998; 98WO-US023826.
 XX
 XX 07-NOV-1997; 97US-0064761P.
 XX
 XX (BIOJ) BIOGEN INC.
 XX
 XX Browning J;
 XX
 XX WPI; 1999-418423/35.
 XX N-PSDB; AAX59349.
 XX
 PT Novel B-cell myelin oligodendrocyte glycoproteins.
 XX
 XX Claim 2; Page 43; 43pp; English.
 XX
 CC This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)

Query Match 29.4%; Score 60; DB 5; Length 1799;
 Sequence 201 AA;

QY 1 STVYQKCHCHMGTHCHSSDGR-----GVIPERCP 33
 DB 158 STVYQK-----CLTWKGRRLPAVVPAPLPP 186

RESULT 9
 ID AAM50359 standard; protein; 1799 AA.
 AC AAM50359;
 XX
 XX 18-FEB-2002 (first entry) *
 XX
 DE Mouse laminin-15 beta 2 chain.
 XX
 KW Laminin-15; mouse; retina; eye; therapy; ophthalmological;
 KW antiinflammatory; rod dystrophy; rod-cone dystrophy;
 KW retinitis pigmentosa; macular degeneration; retinal detachment.
 XX
 OS Mus musculus.
 XX
 XX WO200183516-A1.
 XX
 XX 08-NOV-2001.
 XX
 XX 01-MAY-2001; 2001WO-US013943.
 XX
 XX 01-MAY-2000; 2000US-0200863P.
 XX
 XX (MASS-) MASSACHUSETTS GEN HOSPITAL.
 XX
 XX Burgeson RE, Brunken W, Champiaud M, Hunter D;
 XX WPI; 2002-041478/05.
 XX
 XX Novel substantially pure preparation comprising laminin having laminin
 XX chain alpha 5, beta 2, and gamma 3, useful for treating retinal disorders
 XX such as retinitis pigmentosa, macular degeneration, retinal detachment.
 XX
 XX Disclosure; Fig 3A; 58pp; English.
 XX
 CC The present sequence is that of the beta 2 chain of mouse laminin-15, a
 CC novel member of the laminin family that is produced in the retina. The
 CC retina produces 2 novel laminin trimers: laminin-14 (alpha 4, beta 2,
 CC gamma 3) and laminin-15 (alpha 5, beta 2, gamma 3). These are expressed
 CC within the inter-photoreceptor matrix and in the outer plexiform layer,
 CC and may serve to stabilise retinal synapses. The invention provides
 CC laminin-15 preparations and cells comprising a nucleic acid encoding the
 CC laminin alpha 5, beta 2 and gamma 3 chains, and which are capable of
 CC producing laminin-15. The laminin-15 preparation is used in claimed
 CC methods of: increasing retina immunophoreceptor matrix stability;
 CC increasing the stability of retina photoreceptor compounds, especially an
 CC outer segment, inner segment or synapse; increasing retina adhesion;
 CC treating a disorder associated with retina degeneration, especially rod
 CC dystrophy, rod-cone dystrophy, retinitis pigmentosa, macular degeneration
 CC and retinal detachment; increasing the stability of synapses of the
 CC central nervous system or peripheral nervous system; stimulating an
 CC neuroregeneration axon outgrowth or synapse formation; preparing an
 CC implant, e.g. a catheter, artificial joint, retinal implant, timed
 CC releasing device, neural cell growth guide or artificial tissue, by
 CC coating with the laminin-15 preparation; and increasing photosensitivity
 CC by implanting a tip coated with the laminin-15 preparation into an eye.
 CC The laminin may be recombinant, and the 3 chains co-expressed in the same
 CC cell or expressed in different cells

Query Match 29.4%; Score 60; DB 5; Length 1799;
 Sequence 1799 AA;

Best Local Similarity 29.8%; Pred. No. 1.7e+02;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

QY 5 YQCKCHCW-----GTHCHSSD-GPRGVIPERC 32
: : : : :
Db 1111 FTGQCHAGAGFGRTSCQELYNWDPQLQCRACDCDPRG-IDXPC 1156

RESULT 10
AAW50895
ID AAW50895 standard; protein; 1801 AA.
XX AC AAW50895;
XX DT 07-DEC-1998 (first entry)
XX DE Rat laminin B2 chain.
XX XX Laminin; rat; beta-amylod; amyloidosis; Alzheimer's disease;
KW Down's syndrome; hereditary cerebral haemorrhage; inflammation;
KW malignancy; Familial Mediterranean Fever; multiple myeloma;
KW type II diabetes; prion disease; Creutzfeldt-Jacob disease; CJD;
KW Gerstmann-Straussler syndrome; kuru; scrapie; haemodialysis;
KW carpal tunnel syndrome; senile cardiac amyloid polynuropathy;
KW Familial Amyloidotic Polynuropathy; thyroid carcinoma; diagnosis;
KW therapy.
XX KW Rattus sp.
XX OS
XX XX W09815179-Al.
PN XX
PD 16-APR-1998.
XX
XX 08-OCT-1997; 97WO-US018145.
XX PF
XX 08-OCT-1996; 96US-0027981P.
PR XX
XX (UNIW) UNIV WASHINGTON.
PA
XX Castillo G, Snow AD;
PI
XX WPI; 1998-240534/21.
DR
XX
PT Use of laminin and fragments - for developing products for use in the
PT diagnosis and treatment of amyloid disease, e.g. Alzheimer's disease or
PT CJD.
XX
PS Claim 15; Page 94-97; 132pp; English.

This is the amino acid sequence of the rat laminin B2 chain. The primary object of the invention is to use laminin, laminin-derived protein fragments and/or laminin-derived polypeptides as potent inhibitors of amyloid formation, deposition, accumulation and/or persistence in Alzheimer's disease and other amyloidoses. The laminin products (see AAW50888-98) may include mammalian laminin A or A1 chain, laminin B1 or B2 chain, laminin A2 chain (merosin), laminin G1 chain, the globular repeats of the laminin A1 chain and the beta-amyloid binding domain of the laminin A chain. A claimed method for treating an amyloid disease comprises administering a polypeptide having a conformational similarity to a fragment of a laminin protein. A method for diagnosing an amyloid disease involves determining levels of laminin in a sample. Production of laminin or its fourth globular repeat in vivo provides a method for in vivo inhibition of beta-amyloid amyloidosis. The products and methods can be used for the diagnosis, prognosis, monitoring and treatment of amyloidoses such as Alzheimer's disease, Down's syndrome and hereditary cerebral haemorrhage with amyloidosis of the Dutch type (where the specific amyloid is the beta-amyloid protein), the amyloidosis associated with chronic inflammation, various forms of malignancy and Familial Mediterranean Fever (AA amyloid or inflammation-associated amyloidosis), the amyloidosis associated with multiple myeloma and other B-cell abnormalities (AL amyloid), the amyloidosis associated with type II diabetes (amylin or islet amyloid), the amyloidosis associated with prion diseases including Creutzfeldt-Jacob disease, Gerstmann-Straussler


```
Query Match      28.9%; Score 59; DB 3; Length 2813;
Best Local Similarity 30.6%; Pred. No. 3.4e+02;
Matches 15: Conservative 4; Mismatches 10; Indels
```

QY 4 YQOK-----CHCMG-THCHSDG-----PRGVIPEPRC 32
Db 710 YDGEIFQPEDIFSDHHTWCYCEDGFHCTTSGGLGSLLENVLSPPRC 758

Search completed: February 26, 2004, 12:10:15
Job time : 25.6828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 11.4668 Seconds
(without alignments)
607.670 Million cell updates/sec

Title: US-10-036-444-6
Perfect score: 204
Sequence: 1 STVYQKCHCHMGTHCHSSDGPGRVPEPRCP 33

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	33	US-10-036-444-6	Sequence 6, Appli
2	204	100.0	190	US-10-036-444-2	Sequence 2, Appli
3	61.5	30.1	636	US-09-796-753-100	Sequence 100, App
4	61.5	30.1	636	US-09-796-753-124	Sequence 124, App
5	60	29.4	1799	US-09-845-583-6	Sequence 6, Appli
6	60	29.4	1801	US-09-938-275-8	Sequence 8, Appli
7	59	28.9	2813	US-09-886-900-2	Sequence 2, Appli
8	58.5	28.7	148	US-10-106-698-5317	Sequence 5317, Ap
9	58.5	28.7	2469	US-10-190-115-2	Sequence 2, Appli
10	58.5	28.7	2469	US-10-369-072-2	Sequence 2, Appli
11	58	28.4	845	US-09-983-531A-6	Sequence 6, Appli
12	56.5	27.7	969	US-10-052-648A-34	Sequence 34, Appli
13	56.5	27.7	969	US-10-052-648A-35	Sequence 35, Appli
14	55.5	27.2	601	US-10-365-227-20	Sequence 20, Appli
15	55.5	27.2	1140	US-10-092-390-2	Sequence 2, Appli

16	55.5	27.2	1140	15	US-10-052-648A-33	Sequence 33, Appli
17	54.5	26.7	580	14	US-10-128-714-3183	Sequence 3183, Ap
18	54.5	26.7	680	14	US-10-128-714-8183	Sequence 8183, Ap
19	54.5	26.7	1535	14	US-10-189-971-14	Sequence 14, Appli
20	54.5	26.7	1570	14	US-10-189-971-12	Sequence 12, Appli
21	54.5	26.7	1593	14	US-10-189-971-4	Sequence 4, Appli
22	54.5	26.7	1628	14	US-10-189-971-2	Sequence 2, Appli
23	54.5	26.7	2710	13	US-10-153-273-12	Sequence 12, Appli
24	54	26.5	497	10	US-09-796-753-116	Sequence 116, App
25	54	26.5	928	15	US-10-052-648A-4	Sequence 4, Appli
26	54	26.5	928	15	US-10-052-648A-6	Sequence 6, Appli
27	54	26.5	1034	15	US-10-052-648A-31	Sequence 31, Appli
28	54	26.5	1034	15	US-10-052-648A-32	Sequence 32, Appli
29	54	26.5	1037	15	US-10-052-648A-8	Sequence 8, Appli
30	54	26.5	1037	15	US-10-052-648A-10	Sequence 10, Appli
31	54	26.5	1050	10	US-09-796-753-114	Sequence 114, App
32	53.5	26.2	2749	15	US-10-360-101-265	Sequence 265, App
33	53	26.0	597	15	US-10-369-493-5394	Sequence 5394, Ap
34	53	26.0	818	15	US-10-436-715-59	Sequence 59, Appli
35	52.5	25.7	577	13	US-10-072-841-29	Sequence 29, Appli
36	52.5	25.7	615	15	US-10-291-172-192	Sequence 192, App
37	52.5	25.7	965	14	US-10-221-097-51	Sequence 51, Appli
38	52.5	25.7	1374	14	US-10-205-823-93	Sequence 93, Appli
39	52	25.5	165	14	US-10-411-224-101	Sequence 101, App
40	52	25.5	165	15	US-10-047-021-101	Sequence 101, App
41	52	25.5	186	14	US-10-140-164-5	Sequence 5, Appli
42	52	25.5	186	14	US-10-073-333A-5	Sequence 5, Appli
43	52	25.5	255	10	US-09-866-050A-692	Sequence 692, App
44	52	25.5	255	10	US-09-852-472-24	Sequence 24, Appli
45	52	25.5	275	9	US-09-790-264-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-10-036-444-6
; Sequence 6, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-6

Query Match 100.0%; Score 204; DB 13; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSSDGPGRVPEPRCP 33
Db 1 STVYQKCHCHMGTHCHSSDGPGRVPEPRCP 33

RESULT 2
US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1

GENERAL INFORMATION:
APPLICANT: INNATE PHARMA S.A.S.
TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"
FILE REFERENCE: SEQ-FR-1060
CURRENT APPLICATION NUMBER: US/10/036,444
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 09/440,514
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: 09/456,199
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 2
LENGTH: 190
TYPE: PRT
ORGANISM: Human NK cell
US-10-036-444-2

Query Match 100.0%; Score 204; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.4e-16;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSDGPRGVIPRCP 33
Db 158 STVYQKCHCHMGTHCHSDGPRGVIPRCP 190

RESULT 3
US-09-796-753-100
Sequence 100, Application US/09796753
Publication No. US20030027998A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-227-999
CURRENT APPLICATION NUMBER: US/09/796,753
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/183,175
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 09/223,094
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/223,546
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/224,246
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/259,388
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/122,458
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 09/312,359
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/336,536
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/342,687
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 09/345,464
PRIOR FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: 09/365,164
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/399,723
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 09/409,634
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 09/471,179
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: 09/474,071
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/474,072
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 09/514,010

PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: 09/516,745
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 09/572,002
PRIOR FILING DATE: 2000-05-14
PRIOR APPLICATION NUMBER: 09/597,993
PRIOR FILING DATE: 2000-06-19
PRIOR APPLICATION NUMBER: 09/599,596
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 09/630,334
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 09/606,565
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/606,317
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 09/665,666
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: 09/677,751
PRIOR FILING DATE: 2000-09-30
NUMBER OF SEQ ID NOS: 162
SEQ ID NO 100
LENGTH: 636
TYPE: PRT
ORGANISM: Rattus sp.
US-09-796-753-100

Query Match 30.1%; Score 61.5; DB 10; Length 636;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSDGPRGVIP---EPRC 32
Db 411 CQCHGATGATGQDSCVCIEGTGPNC 437

RESULT 4
US-09-796-753-124
Sequence 124, Application US/09796753
Publication No. US20030027998A1
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-227-999
CURRENT APPLICATION NUMBER: US/09/796,753
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 09/183,175
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 09/223,094
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/223,546
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/224,246
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 09/259,388
PRIOR FILING DATE: 1999-02-26
PRIOR APPLICATION NUMBER: 60/122,458
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: 09/312,359
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/336,536
PRIOR FILING DATE: 1999-06-18
PRIOR APPLICATION NUMBER: 09/342,687
PRIOR FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: 09/345,464
PRIOR FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: 09/365,164
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 09/399,723
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 09/409,634
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 09/471,179
PRIOR FILING DATE: 1999-12-23

```
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1998-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1998-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 124
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-796-753-124
```

```
Query Match 30.1%; Score 61.5; DB 10; Length 636;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;
```

```
QY 9 CHCHMGTHCHSSDGRGVIP---EPEC 32
Db 411 CQCHGATCHPQDGSVCIPGTGPNC 437
```

```
RESULT 5
US-09-845-583-6
; Sequence 6, Application US/09845583
; Patent No. US20020142954A1
; GENERAL INFORMATION:
; APPLICANT: Burgeson, Robert
; APPLICANT: Brunken, William Joseph
; APPLICANT: Champlaud, Marie-France
; APPLICANT: Hunter, Dale
; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
; FILE REFERENCE: 10287-056001
; CURRENT APPLICATION NUMBER: US/09/845,583
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/200,863
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1799
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-845-583-6
```

```
Query Match 29.4%; Score 60; DB 9; Length 1799;
Best Local Similarity 29.8%; Pred. No. 1.4e+02;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;
```

```
QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32
Db 1111 FTQCHCHAGFGGTCSEQELWGDPLQCRACDCDPRG-IDKPCQ 1156
```

```
RESULT 6
```

```
US-09-938-275-8
; Sequence 8, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Caettillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; TITLE OF INVENTION: of Laminin and Laminin-Derived Protein Fragments
; FILE REFERENCE: PROTEO.P03
; CURRENT APPLICATION NUMBER: US/09/938,275
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1801
; TYPE: PRT
; ORGANISM: Rattus No. US20020111309A1vegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Swissprot P15800
; DATABASE ENTRY DATE: 1990-04-01
US-09-938-275-8
```

```
Query Match 29.4%; Score 60; DB 9; Length 1801;
Best Local Similarity 29.8%; Pred. No. 1.4e+02;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;
```

```
QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32
Db 1113 FTQCHCHAGFGGTCSEQELWGDPLQCRACDCDPRG-IDKPCQ 1158
```

```
RESULT 7
US-09-886-900-2
; Sequence 2, Application US/09886900
; Patent No. US20020137051A1
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; Yuzbasiyan-Gurkan, Wilma
; Schall, William D
; Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patenlin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/886,900
; FILING DATE: 21-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/896,449
; FILING DATE: 18-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2813 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-886-900-2

Query Match      28.9%; Score 59; DB 9; Length 2813;
Best Local Similarity 30.6%; Pred. No. 2.8e+02;
Matches 15; Conservative 4; Mismatches 10; Indels 20; Gaps 3;

QY 4 YVQK-----CHCHMG-THCHSDG-----PRGVIPPRC 32
Db 710 YVDGEIFQPEDIFSDHHHTWCYCEDGFHCTTSGGLSLPFPVLSPPRC 758

RESULT 8
US-10-106-698-5317
; Sequence 5317, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 5317
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5317

Query Match      28.7%; Score 58.5; DB 14; Length 148;
Best Local Similarity 43.5%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 10 HCHMGTHCHSDGPRGVIPPRC 32
Db 83 HCHLGYRCHRPQREGI---PRC 102

RESULT 9
US-10-190-115-2
; Sequence 2, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Casman, William M.
; APPLICANT: Grosse, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Spaderna, Steven K.

Query Match      28.7%; Score 58.5; DB 15; Length 2469;
Best Local Similarity 42.3%; Pred. No. 2.8e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHCH-HSSDGPGRGVIPPR 31
Db 1345 GQVKCRKGEOCVHTASGRCFCPSPR 1370

RESULT 10
US-10-369-072-2
; Sequence 2, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spyttek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
```

```
; APPLICANT: Spyttek, Kimberly A.
; APPLICANT: Szerkeres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 2
; LENGTH: 2469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-2

Query Match      28.7%; Score 58.5; DB 15; Length 2469;
Best Local Similarity 42.3%; Pred. No. 2.8e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHCH-HSSDGPGRGVIPPR 31
Db 1345 GQVKCRKGEOCVHTASGRCFCPSPR 1370

RESULT 10
US-10-369-072-2
; Sequence 2, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spyttek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
```

```
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 2469
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-2

Query Match      28.7%; Score 58.5; DB 15; Length 2469;
Best Local Similarity 42.3%; Pred. No. 2.8e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMTHC-HSSDGRGVIPER 31
DB 1345 GQVKCKGECQVHTASGRFCPCSPR 1370

RESULT 11
US-09-983-531A-6
/ Sequence 6, Application US/09983531A
/ Patent No. US20020147132A1
/ GENERAL INFORMATION:
/ APPLICANT: Fujisawa, Atsuko
/ APPLICANT: Yamakawa, Toru
/ APPLICANT: Shirakawa, Kamon
/ APPLICANT: Chitose, Ozii
/ APPLICANT: Ogawa, Naoki
/ TITLE OF INVENTION: Meltrins
/ FILE REFERENCE: 11-22-99 sequence submission
/ CURRENT APPLICATION NUMBER: US/09/983,531A
/ CURRENT FILING DATE: 2001-10-24
/ PRIOR APPLICATION NUMBER: JP 8-61756
/ PRIOR FILING DATE: 1996-02-23
/ PRIOR APPLICATION NUMBER: PCT/JP96/03017
/ PRIOR FILING DATE: 1996-10-17
/ NUMBER OF SEQ ID NOS: 44
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 845
/ TYPE: PRT
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism: Clone,
/ OTHER INFORMATION: JM109(pBSmel-gamma), mouse meltrin gamma
US-09-983-531A-6

Query Match      28.4%; Score 58; DB 9; Length 845;
Best Local Similarity 47.6%; Pred. No. 1.2e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 4; Gaps 1;

QY 6 GQKCH-----CHMGTHCHSSDG 22
```

```
DB 647 QGKCHGHGVCNKNKNCEDG 667

RESULT 12
US-10-052-648A-34
/ Sequence 34, Application US/10052648A
/ Publication No. US20040005558A1
/ GENERAL INFORMATION:
/ APPLICANT: Anderson, David
/ APPLICANT: Burgess, Catherine
/ APPLICANT: Casman, Stacie
/ APPLICANT: Colman, Steven
/ APPLICANT: Edinger, Shlomit R.
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Gunther, Erik
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: MacDougall, John R.
/ APPLICANT: Mehraban, Fuad
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Rothenberg, Mark
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Smithson, Glennda
/ APPLICANT: Spytek, Kimberly A.
/ APPLICANT: Stone, David J.
/ APPLICANT: Vernhet, Corine A.M.
/ APPLICANT: Zerhusen, Bryan D.
/ TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
/ TITLE OF INVENTION: USING THE SAME
/ FILE REFERENCE: 21402-250 (CURA-550)
/ CURRENT APPLICATION NUMBER: US/10/052,648A
/ CURRENT FILING DATE: 2002-12-09
/ PRIOR APPLICATION NUMBER: 60/262,454
/ PRIOR FILING DATE: 2001-01-18
/ PRIOR APPLICATION NUMBER: 60/272,920
/ PRIOR FILING DATE: 2001-03-02
/ PRIOR APPLICATION NUMBER: 60/284,549
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/303,229
/ PRIOR FILING DATE: 2001-07-05
/ PRIOR APPLICATION NUMBER: 60/262,892
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: 60/263,605
/ PRIOR FILING DATE: 2001-01-23
/ PRIOR APPLICATION NUMBER: 60/269,098
/ PRIOR FILING DATE: 2001-02-15
/ PRIOR APPLICATION NUMBER: 60/264,159
/ PRIOR FILING DATE: 2001-01-25
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/271,855
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 97
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 34
/ LENGTH: 969
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-052-648A-34

Query Match      27.7%; Score 56.5; DB 15; Length 969;
Best Local Similarity 34.4%; Pred. No. 2e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 7; Gaps 1;

QY 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33
DB 635 CSCHNGASCASDGDGACTPGWTGLFCTQRCP 666

RESULT 13
US-10-052-648A-35
```

; Sequence 35, Application US/10052648A
; Publication No. US2004005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catharine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (848)...(889)
; OTHER INFORMATION: Where Xaa is any amino acid
US-10-052-648A-35

Query Match 27.7%; Score 56.5; DB 15; Length 969;
Best Local Similarity 34.4%; Pred. No. 2e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33
Db 635 CSCENGASCAEDGACHCTPGWTGLFCTQRCP 666

RESULT 14
US-10-365-227-20
; Sequence 20, Application US/10365227
; Publication No. US20030143632A1

; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-323001
; CURRENT APPLICATION NUMBER: US/10/365,227
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US/09/802,582
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-365-227-20

Query Match 27.2%; Score 55.5; DB 14; Length 601;
Best Local Similarity 34.4%; Pred. No. 1.6e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33
Db 311 CNCHNGAFCSAYDGECKCTPGWTGLYCTQRCP 342

RESULT 15
US-10-092-390-2
; Sequence 2, Application US/10092390
; Publication No. US20030013865A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; TITLE OF INVENTION: No. US20030013865A1 Human EGF-Family Proteins and Polynucleoti

; FILE REFERENCE: LEX-0317-USA
; CURRENT APPLICATION NUMBER: US/10/092,390
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/275,013
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1140
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-092-390-2

Query Match 27.2%; Score 55.5; DB 14; Length 1140;
Best Local Similarity 34.4%; Pred. No. 3e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33
Db 716 CNCHNGAFCSAYDGECKCTPGWTGLYCTQRCP 747

Thu Feb 26 12:38:28 2004

us-10-036-444-6.rapb

Page 7

Search completed: February 26, 2004, 12:33:57
Job time : 13.4668 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 6.38992 Seconds
(without alignments)
266.616 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMTHCHSSDGRGVIPPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pep:*
- 5: /cgn2_6/prodata/2/iaa/PCTUS COMB.pep:*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	60	29.4	1799	4	US-09-845-583A-6
2	59	28.9	2813	3	US-08-896-449A-2
3	59	28.9	2813	3	US-09-132-652-2
4	58.5	28.7	1015	1	US-08-537-210A-1
5	58.5	28.7	1015	3	US-09-113-825-1
6	58.5	28.7	2471	1	US-08-185-432-16
7	58.5	28.7	2471	1	US-08-083-590A-19
8	58.5	28.7	2471	3	US-08-532-384-19
9	58.5	28.7	2471	4	US-08-899-232-1
10	54.5	26.7	2710	2	US-08-568-459A-12
11	54.5	26.7	2710	2	US-08-487-826B-12
12	54.5	26.7	2710	2	US-09-210-288-12
13	54.5	26.7	3060	4	US-08-487-826B-14
14	52.5	25.7	577	2	US-07-728-215-29
15	52.5	25.7	577	4	US-08-938-085A-29
16	52.5	25.7	577	4	US-10-072-844-29
17	52.5	25.7	577	4	US-10-072-838-29
18	52.5	25.7	577	4	US-10-072-841A-29
19	52.5	25.7	799	1	US-08-054-077C-2
20	52	25.5	278	4	US-09-724-864-52
21	52	25.5	420	4	US-09-252-991A-27434
22	52	25.5	470	4	US-09-252-991A-27804
23	51	25.0	70	4	US-09-381-546-21
24	51	25.0	207	3	US-08-974-022-47
25	51	25.0	207	3	US-08-795-445A-47
26	51	25.0	207	3	US-08-795-447A-47
27	51	25.0	207	3	US-08-974-186-47

Sequence 47, Appl
Sequence 133, App
Sequence 47, Appl
Sequence 2, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 24401, A
Sequence 5, Appl
Sequence 5, Appl
Sequence 8, Appl
Sequence 11, Appl
Sequence 132, App
Sequence 38, Appl
Sequence 38, Appl
Sequence 325, App

28 51 25.0 207 3 US-08-795-446B-47
29 51 25.0 207 4 US-08-706-945D-133
30 51 25.0 207 4 US-08-577-788C-47
31 51 25.0 325 1 US-08-292-549-2
32 51 25.0 325 3 US-09-042-785A-9
33 51 25.0 325 5 PCT-US91-02207-2
34 50.5 24.8 298 3 US-09-080-044-9
35 50.5 24.8 298 4 US-09-531-857A-9
36 50.5 24.8 771 1 US-09-252-991A-24401
37 50 24.5 277 4 US-08-690-457-5
38 50 24.5 277 2 US-08-628-187-5
39 50 24.5 277 3 US-08-432-071-3
40 50 24.5 1798 4 US-09-845-583A-8
41 50 24.5 1798 4 US-09-561-709B-11
42 49.5 24.3 299 3 US-09-188-930-192
43 49 24.0 115 3 US-08-513-974B-38
44 49 24.0 115 4 US-09-461-436B-38
45 49 24.0 132 3 US-08-513-974B-325

ALIGNMENTS

RESULT 1

US-09-845-583A-6
; Sequence 6, Application US/09845583A
; Patent No. 6635616

GENERAL INFORMATION:

; APPLICANT: Burgeson, Robert
; APPLICANT: Brunken, William Joseph
; APPLICANT: Champlaud, Marie-France
; APPLICANT: Hunter, Dale
; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
; FILE REFERENCE: 10287-056001
; CURRENT APPLICATION NUMBER: US/09/845,583A
; PRIOR FILING DATE: 2001-04-30
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1799
; TYPE: PRT
; ORGANISM: Mus musculus

Query Match

Query Match 29.4%; Score 60; DB 4; Length 1799;
Best Local Similarity 29.8%; Pred. No. 25;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

Cy 5 YQKCHCHM-----GTHCHSSD-GPRGVIPPRC 32

Db 1111 FTGQCHCHAGFGGRTCTCQELYWGDPGLQCRACDCPRG-IDKQC 1156

RESULT 2

US-08-896-449A-2
; Sequence 2, Application US/08896449A
; Patent No. 6040143

GENERAL INFORMATION:

; APPLICANT: Ventr, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; TITLE OF INVENTION: FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCE: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA

ZIP: 48098
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,449A
FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
TELEX: 287637
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2813 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-896-449A-2

Query Match 28.9%; Score 59; DB 3; Length 2813;
Best Local Similarity 30.6%; Pred. No. 52;
Matches 15; Conservative 4; Mismatches 10; Indels 20; Gaps 3;

Qy 4 YQOK-----CHCMG-THCHSSDG-----PRGVIPPRC 32
Db 710 YDGEIPQPEDIFSDHHTWCYCDGFMHCTTSGGLGSLLPNPNVLSPPRC 758

RESULT 3

US-09-132-652-2
Sequence 2, Application US/09132652
Patent No. 6074832
GENERAL INFORMATION:
APPLICANT: Venta, Patrick J
APPLICANT: Yuzbasiyan-Gurkan, Vilma
APPLICANT: Schall, William D
APPLICANT: Brewer, George J
APPLICANT: Duffendeck, John
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS
FILE REFERENCE: 2115S-001226CPB
CURRENT APPLICATION NUMBER: US/09/132,652
CURRENT FILING DATE: 1998-08-11
EARLIER APPLICATION NUMBER: 08/896,449
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 2813
TYPE: PRT
ORGANISM: Canis familiaris
US-09-132-652-2

Query Match 28.9%; Score 59; DB 3; Length 2813;
Best Local Similarity 30.6%; Pred. No. 52;
Matches 15; Conservative 4; Mismatches 10; Indels 20; Gaps 3;

Qy 4 YQOK-----CHCMG-THCHSSDG-----PRGVIPPRC 32
Db 710 YDGEIPQPEDIFSDHHTWCYCDGFMHCTTSGGLGSLLPNPNVLSPPRC 758

RESULT 4

US-08-537-210A-1
Sequence 1, Application US/08537210A
Patent No. 5780300
GENERAL INFORMATION:

APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Fortini, Mark
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA: US/08/537,210A
APPLICATION NUMBER: US/08/537,210A
FILING DATE: 29-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1015 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: hum N (Human No. 5780300ch 2)
LOCATION: 1155..2169
OTHER INFORMATION: Highly conserved ankyrin repeat
OTHER INFORMATION: region of No. 5780300ch
US-08-537-210A-1

Query Match 28.7%; Score 58.5; DB 1; Length 1015;
Best Local Similarity 42.3%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

Qy 7 GKCHCMGTHC-HSSDGRGVIPPR 31
Db 193 GQVCKRGECVHTASGRFCPSPR 218

RESULT 5

US-09-113-825-1
Sequence 1, Application US/09113825
Patent No. 6149902
GENERAL INFORMATION:

APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Fortini, Mark
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

```

;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,825
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/537,210
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1015 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: hum N (Human No. 6149902ch 2)
; LOCATION: 1155...2169
; OTHER INFORMATION: Highly conserved ankyrin repeat
; OTHER INFORMATION: region of No. 6149902ch
US-09-113-825-1

Query Match 28.7%; Score 58.5; DB 3; Length 1015;
Best Local Similarity 42.3%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVIPER 31
Db 193 GQVKCKGECQCVHTASGPRCFCPSR 218

RESULT 6
US-08-185-432-16
; Sequence 16, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Atavanis-Teakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Jian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: E30
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872

```

```

;
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-16

Query Match 28.7%; Score 58.5; DB 1; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVIPER 31
Db 1347 GQVKCKGECQCVHTASGPRCFCPSR 1372

RESULT 7
US-08-083-590A-19
; Sequence 19, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Teakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-083-590A-19

Query Match 28.7%; Score 58.5; DB 1; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVIPER 31
Db 1347 GQVKCKGECQCVHTASGPRCFCPSR 1372

```

RESULT 8
US-08-532-384-19
; Sequence 19, Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,384
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,590
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-532-384-19

Query Match 28.7%; Score 58.5; DB 3; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVPEPR 31
DB 1347 GQVKCRKGEQCVHTASGRFCFSPR 1372

RESULT 9
US-08-899-232-1
; Sequence 1, Application US/08899232
; Patent No. 6436650
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Qi, Huilin
; TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
; FILE REFERENCE: 7326-046
; CURRENT APPLICATION NUMBER: US/08/899,232
; CURRENT FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2471
; TYPE: PRT
; ORGANISM: Homo sapiens

US-08-899-232-1

Query Match 28.7%; Score 58.5; DB 4; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVPEPR 31
DB 1347 GQVKCRKGEQCVHTASGRFCFSPR 1372

RESULT 10
US-08-568-459A-12
; Sequence 12, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellem, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 820 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2710 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium falciparum
US-08-568-459A-12

Query Match 26.7%; Score 54.5; DB 2; Length 2710;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKCHCHMGTHCHS-SDGPRGVPEPRCP 33
DB 1700 YPEKDCYQKGVHPSIPPPPPVQPOPEAP 1729

RESULT 11
US-08-487-826B-12
; Sequence 12, Application US/08487826B
; Patent No. 5993827

GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-08-487-826B-12

Query Match 26.7%; Score 54.5; DB 2; Length 2710;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKQKCHCMGTHCHS-SDGPRGVIPRCP 33
DB 1700 YPEKDCYQGHVPSIPPPPPVQPOPEAP 1729

RESULT 12
US-09-210-288-12
Sequence 12, Application US/09210288
Patent No. 632026
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California

COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,288
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Fuller, Michael
REGISTRATION NUMBER: 36,516
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2710 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium falciparum
US-09-210-288-12

Query Match 26.7%; Score 54.5; DB 4; Length 2710;
Best Local Similarity 40.0%; Pred. No. 1.8e+02;
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKQKCHCMGTHCHS-SDGPRGVIPRCP 33
DB 1700 YPEKDCYQGHVPSIPPPPPVQPOPEAP 1729

RESULT 13
US-08-487-826B-14
Sequence 14, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israel, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CPI

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3060 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-487-826B-14

Query Match 26.7%; Score 54.5; DB 2; Length 3060;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKQKCHCMGTHCHS-SDGPRGVIPPRCP 33
DB 1698 YPEKDCYQKGVPSIPPPPPVQPOPEAP 1727

RESULT 14

US-07-728-215-29
Sequence 29, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 5962643el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 577 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-728-215-29

Query Match 25.7%; Score 52.5; DB 2; Length 577;
Best Local Similarity 34.2%; Pred. No. 68;
Matches 13; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

QY 4 YQKQKCHC---HMGTHCH-----SSDGRGVIPPRCP 32
DB 333 YQCGVCACNPGHMGPHCECGEDTLSTDCKETPDHPC 370

RESULT 15

US-08-938-085A-29
Sequence 29, Application US/08938085A
Patent No. 6339148
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
APPLICANT: Pytela, Robert
TITLE OF INVENTION: A No. 6339148el Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,085A
FILING DATE: 26-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,215
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 023070-080210US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 577 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-085A-29

Query Match 25.7%; Score 52.5; DB 4; Length 577;
Best Local Similarity 34.2%; Pred. No. 68;
Matches 13; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

QY 4 YQKQKCHC---HMGTHCH-----SSDGRGVIPPRCP 32
DB 333 YQCGVCACNPGHMGPHCECGEDTLSTDCKETPDHPC 370

Search completed: February 26, 2004, 12:16:18
Job time : 7.38992 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 2.2679 Seconds
(without alignments)
636.214 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPEIR-TLRGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.5	54.7	155	2 F71637	peptidoglycan-asso
2	46.5	54.7	155	2 H97849	peptidoglycan-asso
3	45	52.9	353	2 T01256	probable anthocyan
4	43	50.6	305	2 T36261	hypothetical prote
5	43	50.6	514	2 A49838	site-specific reco
6	43	50.6	514	2 AF1988	fdxN element site-
7	43	50.6	687	2 AD2347	ribonuclease E lim
8	42	49.4	101	2 G75512	conserved hypothet
9	42	49.4	372	1 S28296	hypothetical prote
10	42	49.4	478	1 S60754	transcription fact
11	42	49.4	486	1 TVHUEG	transforming prote
12	42	49.4	535	2 S75159	hypothetical prote
13	42	49.4	1317	2 F83310	conserved hypothet
14	41.5	48.8	60	2 S68769	short neurotoxin -
15	41	48.2	272	2 S35641	flagellar motor ap
16	41	48.2	347	2 S71436	farnesyl-pyrophosp
17	41	48.2	368	2 S67507	morphogen lefty pr
18	41	48.2	517	2 C89840	conserved hypothet
19	41	48.2	597	2 T08681	adenosylhomocystei
20	41	48.2	788	2 AG0786	secreted effector
21	41	48.2	805	2 T24399	hypothetical prote
22	40	47.1	196	2 H64968	acetyl CoA acetyl
23	40	47.1	209	2 B83329	probable acyltrans
24	40	47.1	316	2 E75566	probable repressor
25	40	47.1	341	2 S42699	hypothetical prote
26	40	47.1	341	2 B35769	cAMP response elem
27	40	47.1	341	2 S03343	CAMP response elem
28	40	47.1	349	2 D75419	proline dipeptidas
29	40	47.1	396	2 E64049	cystathionine beta

30	40	47.1	443	2 H96841	hypothetical prote
31	40	47.1	664	2 T48258	kinesin-like prote
32	40	47.1	690	2 C69208	DNA helicase relat
33	40	47.1	3707	2 S18262	heparan sulfate pr
34	39.5	46.5	1694	2 S50865	sialoadhesin - mou
35	39	45.9	107	2 F81747	hypothetical prote
36	39	45.9	108	2 T13133	hypothetical prote
37	39	45.9	293	2 S64925	protein gp46 - pha
38	39	45.9	371	2 AH1263	hypothetical prote
39	39	45.9	434	2 F41880	probable tRNA (5-m
40	39	45.9	515	2 E71447	47.87K diazotroph
41	39	45.9	527	2 S46888	probable PSII D1 p
42	39	45.9	916	2 G75417	hypothetical prote
43	39	45.9	1197	2 A91035	SNF2/Rad54 helicase
44	39	45.9	1197	2 C85879	probable sensor fo
45	38.5	45.3	185	2 AB2241	hypothetical prote

ALIGNMENTS

RESULT 1

F71637

peptidoglycan-associated lipoprotein precursor (pal) RP771 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000

C:Accession: F71637

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9823893

A:Accession: F71637

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-155 <AND>

A:Cross-references: GB:AJ235273; GB:AJ235269; NID:G3861237; PIDN:CAAL1598.1; PID:G38612

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: pal; RP771

C:Superfamily: outer membrane protein A

Query Match 54.7%; Score 46.5; DB 2; Length 155;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qv 1 WVSQPPPEIR-TLRGSC 15

Db 74 WLSKHPVKVTVEGHC 89

RESULT 2

H97849

peptidoglycan-associated lipoprotein precursor [imported] - Rickettsia conorii (strain
C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 03-Jun-2002

C:Accession: H97849

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: H97849

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <KOR>

A:Cross-references: GB:AE006914; PIDN:AAL03738.1; PID:G15620330; GSPDB:GN00173

C:Genetics:

A:Gene: pal

C:Superfamily: outer membrane protein A

Query Match 54.7%; Score 46.5; DB 2; Length 155;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 1 WVSQPPEIR-TLEGSC 15
| : | : | : | : |
Db 74 WLSKHPEVKVTVEGHC 89

RESULT 3

T01256

Probable anthocyanidin synthase [imported] - Arabidopsis thaliana
N;Alternate names: SKGL protein homolog F16M14.17
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C;Accession: T01256; F84802
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, A.;Submitted to: The EMBL Data Library, July 1998
A;Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.
A;Reference number: Z14213
A;Accession: T01256
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-353 <ROU>
A;Cross-references: EMBL:AC003028; NID:G3335356; PTD:G3353572
A;Experimental source: cultivar: Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varaken, S.E.; Umayam, L.; Tallon, L.;
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197

A:Status

A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:Cross-references: GB:AE002093; NID:g3335372; PIDD:AAC27173.1; GSPDB:GN00139
C:Gene: F16M14.17; At2g38240
A:Map position: 2
A:Introns: 160/2; 271/3
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

```
Query Match          52.9%; Score:45; DB 2; Length 353;
Best Local Similarity 72.7%; Pred. No.10;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 WVSQPPEIRTL 11
          |||||:|
Db     152 WPSOPPKIREL 163

```

RESULT 4

hypothetical protein SCE68.11 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T32661
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21576
A:Accession: T32661
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-305 <MUR>
A:Cross-references: EMBL:AL079345; PIDN:CAB45347.1; GSPDB:GN00070; SCOEDB:SCE68.11
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE68.11

```
Query Match          50.6%; Score 43; DB 2; Length 305;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 WVSQPEIRTL 11
          |||:|:|
Db     140 WVGRRPREVRTL 15
```

RESULT 5

A49338
site-specific recombinase XisF - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999
C:Accession: A49338
R:Carasco, C.D.; Ramaswamy, K.S.; Ramasubramanian, T.S.; Golden, J.W.
Genes Dev. 8, 74-83, 1994
A:Title: Anabaena xisF gene encodes a developmentally regulated site-specific recombinase
A:Reference number: A49338; MUID:94116860; PMID:8288129
A:Accession: A49338
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-514 <CAR>
A:Cross-references: GB:I23220; NID:G349678; PIDN:AAAL6762.1; PID:G349679
C:Genetics:
A:Gene: xisF

```
Query Match      50.6%; Score 43; DB 2; Length 514;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

Qy	2	VSQPPEIR	LEGS	14
			: :	
Db	414	VEEPPVK	TLRAS	42

RECEIVED

RES0011 6
 AF1988
 fdxN element site-specific recombinase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 C:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AF1988
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch,
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AF1988
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-514 <R>
 A:Cross-references: GB:BA000019; PIDN:BAE73414.1; PID:gl7130804; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: xisF

Query Match 50.6%; Score 43; DB 2; Length 514;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VSPPEIRTEGS 14
| : | | : | |
Db 414 VEEPPEVKTLRAS 42

RECIT.T 7

ribonuclease E [imported] - Nostoc sp. (strain PCC 7120)
AD2347
C:Species: Nostoc sp. PCC 7120
C:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AD2347
R:Kanakako, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch,
Nakazaki, N.; Shimpso, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2347
A>Status: preliminary

F:320-398/Domain: ets DNA-binding domain homology <ETS>

Query Match 49.4%; Score 42; DB 1; Length 486;
Best Local Similarity 46.7%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WVSQPPPIRTLEGGC 15
|:|||||:|:
Db 63 WLSQPPPARVITIMEC 77

RESULT 12

S75159
hypothetical protein slr1771 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75159
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-535 <KAN>
A:Cross-references: EMBL:D90903; GB:AB001339; NID:G1652127; PIDN:BAAL1073.1; PID:d101780
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 49.4%; Score 42; DB 2; Length 535;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPE 7
|:|||||:
Db 203 WISQPE 209

RESULT 13

F83310
conserved hypothetical protein PA2694 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83310
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83310
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1317 <STO>
A:Cross-references: GB:AE004696; GB:AE004091; NID:G9948750; PIDN:AAG06072.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2684
C:Superfamily: rshF protein

Query Match 49.4%; Score 42; DB 2; Length 1317;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WVSQPPPIRTLEGS 14
|:|||||:
Db 155 WAAVPAELQEQGS 168

RESULT 14

S68769

short neurotoxin - black-banded coral snake
N:Alternate names: alpha-neurotoxin

C:Species: *Micurus nigrocinctus*
C:Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 18-Jul-2001
C:Accession: S68769
R:Rosso, J.P.; Vargas-Rosso, O.; Gutierrez, J.M.; Rochat, H.; Bougis, P.E.
Eur. J. Biochem. 238, 231-239, 1996
A:Title: Characterization of alpha-neurotoxin and phospholipase A(2) activities from *Micurus nigrocinctus* in from *Micurus nigrocinctus* nigrocinctus.
A:Reference number: S68769; MUID:96248443; PMID:8665942
A:Accession: S68769

A:Molecule type: protein
A:Residues: 1-60 <ROS>
A:Experimental source: *Micurus nigrocinctus* (black-banded coral snake)
C:Superfamily: snake toxin
C:Keywords: neurotoxin; venom
F:3-22,17-39,41-52,53-58/Disulfide bonds: #status predicted

Query Match 48.8%; Score 41.5; DB 2; Length 60;
Best Local Similarity 64.3%; Pred. No. 6.3;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 3 SQPPPIRTL-EGSC 15
|:|||||:
Db 9 SQPPPIKTCEGQC 22

RESULT 15

S39641
flagellar motor apparatus homolog ytxD - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S39641; B70003
R:Grundy, F.J.; Waters, D.A.; Takova, T.Y.; Henkin, T.M.
Mol. Microbiol. 10, 259-271, 1993
A:Title: Identification of genes involved in utilization of acetate and acetoin in *Bacillus subtilis*
A:Reference number: S39641; MUID:95020526; PMID:7934817
A:Accession: S39641
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <GRU>
A:Cross-references: GB:I117309; NID:9861173; PIDN:AA68282.1; PID:9348048
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berke C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Mauee Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetell Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A.; Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:198044033; PMID:9384377
A:Accession: B70003
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-272 <KUN>
A:Cross-references: GB:299119; GB:AL009126; NID:92635411; PIDN:CAB14951.1; PID:92635457
A:Experimental source: strain 168
C:Genetics:
A:Gene: ytxD
C:Superfamily: flagellar motor rotation protein
C:Keywords: transmembrane protein

Query Match 48.2%; Score 41; DB 2; Length 272;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WVSOPPEIRTEGS 14
Db 255 WQOPKOVTKKGS 268

Search completed: February 26, 2004, 12:14:59
Job time : 5.2679 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 1.27321 Seconds
(without alignments)
613.452 Million cell updates/sec

Title: US-10-036-444-7
Perfect score: 85
Sequence: 1 WVSQPPHRTLEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	51.8	688	1 EOMD_MOUSE	O54839 mus musculus
2	43	50.6	692	1 EOMD_XENLA	P79944 xenopus lae
3	42	49.4	372	1 YLFI_CABEL	Q03571 caenorhabdi
4	42	49.4	383	1 TRMU_PASMU	Q96143 pasteurilla
5	42	49.4	462	1 ERG_HUMAN	P11308 homo sapien
6	42	49.4	478	1 ERG_CHICK	Q90837 gallus gall
7	41.5	48.8	60	1 NX51_MICNI	P80548 micrurus ni
8	41	48.2	272	1 YTXD_BACSU	P39063 bacillus su
9	41	48.2	347	1 FPFS_NEUCR	Q92250 neurospora
10	41	48.2	368	1 LFTB_MOUSE	P57785 mus musculus
11	41	48.2	368	1 TGF4_MOUSE	Q64280 mus musculus
12	41	48.2	686	1 EOMD_HUMAN	Q95936 homo sapien
13	41	48.2	1411	1 Y297_HUMAN	O15040 homo sapien
14	40	47.1	196	1 WBJT_ECOLI	P37750 escherichia
15	40	47.1	341	1 CREB_HUMAN	P16220 homo sapien
16	40	47.1	341	1 CREB_MOUSE	Q01147 mus musculus
17	40	47.1	341	1 CREB_RAT	P15337 rattus norv
18	40	47.1	396	1 METC_HAEIN	P44527 haemophilus
19	40	47.1	690	1 HELS_METTH	O26901 methanobact
20	40	47.1	3707	1 PGBM_MOUSE	Q05793 mus musculus
21	39.5	46.5	1694	1 SN_MOUSE	Q62230 mus musculus
22	39.5	46.5	1709	1 SN_HUMAN	Q95222 homo sapien
23	39	45.9	371	1 TRMU_LISMO	Q8Y114 listeria mo
24	39	45.9	434	1 RP54_RHOSH	Q01194 rhodobacter
25	39	45.9	527	1 YB64_YEAST	P38314 saccharomyc
26	39	45.9	942	1 M3KE_MOUSE	Q9WU16 mus musculus
27	39	45.9	1017	1 TRL3_HUMAN	Q9HCF6 homo sapien
28	39	45.9	1197	1 EVGS_ECO57	P58402 escherichia
29	38	44.7	187	1 RK24_COBAC	Q02764 nicotiana t
30	38	44.7	334	1 GTR8_BOVIN	P58354 bos taurus
31	38	44.7	369	1 PROB_NEIMA	Q9JUK7 neisseria m
32	38	44.7	369	1 PROB_NEIMB	Q9JZ92 neisseria m
33	38	44.7	409	1 LIS1_BOVIN	P43033 bos taurus

RESULT 1

ID	EOMD_MOUSE	STANDARD	PRT	688 AA
AC	O54839; Q9QY07;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Eomesodermin homolog.			
GN	EOMES OR TBR2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=99337662; PubMed=10407135;			
RA	Kimura N., Nakashima K., Ueno M., Taga T.;			
RT	"A novel mammalian T-box-containing gene, Tbr2, expressed in mouse developing brain."			
RL	Brain Res. Dev. Brain Res. 115:183-193 (1999).			
RN	[2]			
RP	SEQUENCE OF 278-457 FROM N.A.			
RX	MEDLINE=98163742; PubMed=9503012;			
RA	Wattler S., Russ A., Evans M., Nehls M.;			
RT	"A combined analysis of genomic and primary protein structure defines the phylogenetic relationship of new members of the T-box family.";			
RL	Genomics 48:24-33 (1998).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	-1- SIMILARITY: Contains 1 T-box domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; AB031037; BAA83416.1; --			
DR	EMBL; AF013281; AAC16233.1; --			
DR	HSPB; P24781; 1XBR.			
DR	MGD; MGI:1201683; Eomes.			
DR	InterPro; IPR008987; P53-like.			
DR	InterPro; IPR001699; TF_T-box.			
DR	Prints; PF00907; T-box; 1.			
DR	SMART; SM00425; TBOX; 1.			
DR	PROSITE; PS01283; TBOX 1; 1.			
DR	PROSITE; PS01264; TBOX 2; 1.			
DR	PROSITE; PS02522; TBOX 3; 1.			
KW	Developmental protein; Transcription regulation; DNA-binding;			
KW	Nuclear protein.			
FT	DOMAIN 27 41			POLY-GLY.
FT	DNA BIND 278 458			T-BOX.
FT	DOMAIN 383 386			POLY-ASN.

ALIGNMENTS

P43034 homo sapien
P43035 mus musculus
Q16999 caenorhabdi
Q01603 drosophila
Q99558 homo sapien
P17139 caenorhabdi
Q14674 homo sapien
O88799 mus musculus
P20956 xenopus lae
Q9YF53 aeropyrum p
P01422 najja haje a
P25675 najja haje h

LIS1_HUMAN
LIS1_MOUSE
ZPRI_CABEL
PERO_DROME
M3KE_HUMAN
CAL14_CABEL
ESF1_HUMAN
ZAN_MOUSE
HVO1_XENLA
RL19_ASRPE
NXS2_NAJHA
NXS2_NAJHH

38 44.7 409
38 44.7 409
38 44.7 458
38 44.7 690
38 44.7 947
38 44.7 1758
38 44.7 1795
38 44.7 5376
37.5 44.1 136
37.5 44.1 155
37 43.5 61
37 43.5 61

```
SQ SEQUENCE 688 AA; 72638 MW; 197B0B989E920B82B CRC64;
Query Match 51.8%; Score 44; DB 1; Length 688;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGGS 14
DB 620 WETPPSIKSLDSS 633

RESULT 2
EOMD XENLA
ID EOMD XENLA STANDARD; PRT; 692 AA.
AC P79944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Somesodermin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
GN XEN
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RX MEDLINE=97133207; PubMed=8978604;
RT Ryan K., Garrett N., Mitchell A., Gurdon J.B.;
RL Cell 87:989-1000(1996).
CC -!- FUNCTION: Involved in mesoderm differentiation. Activates WNT8,
CC brachyury, CHD and MTX.1 expression.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DEVELOPMENTAL STAGE: First expressed at or just after midblastula
CC transition (stage 8). Maximally expressed at stage 10 as an
CC equatorial mesoderm band, more prominently on the dorsal side
CC and around the invaginating dorsal lip.
CC -!- INDUCTION: By activin.
CC -!- DOMAIN: CONTAINS 13 S-P-X-X REPEATS.
CC -!- SIMILARITY: Contains 1 T-box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75996; AAC60061.1; --
CC HSP; P24781; 1XER.
CC InterPro; IPR008967; P53-like.
CC InterPro; IPR001699; TF_T-box.
CC Pfam; PF00907; T-box; 1.
CC PRINTS; PR00937; TBOX.
CC SMART; SM00425; TBOX; 1.
CC PROSITE; PS01283; TBOX_1; 1.
CC PROSITE; PS01284; TBOX_2; 1.
CC PROSITE; PS0252; TBOX_3; 1.
CC Developmental protein; Transcription regulation; DNA-binding;
CC Nuclear protein; Repeat.
CC NUCLEIC BIND 263 443
FT T-BOX.
SQ SEQUENCE 692 AA; 75943 MW; 9D129A67F6357989 CRC64;
Query Match 50.6%; Score 43; DB 1; Length 692;
Best Local Similarity 34.8%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

QY 1 WVSQPPEIRTLGGS-----GSC 15
DB 626 WETPPSIKSLDSSGVYTGC 648

SQ SEQUENCE 372 AA; 42393 MW; 8016734389D97E18 CRC64;
Query Match 49.4%; Score 42; DB 1; Length 372;
Best Local Similarity 61.5%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSQPPEIRTLGGS 14
DB 21 VSAPPEVKRLDGS 33

RESULT 4
TRMU PASMU
ID TRMU PASMU STANDARD; PRT; 383 AA.
AC Q9C1A3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
 DE (SC 2.1.1.61).
 DE TRM OR PM1336.
 OS Pasteurella multocida.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OS Pasteurellaceae; Pasteurella.
 OC NCBI_TaxID=747;
 OX [1]
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc Natl Acad Sci U S A. 98:3460-3465(2001).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing 5-methylaminomethyl-2-
 CC thiouridylate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the trmU family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE006172; AK03420.1;
 CC HAWAP; MF_00144; -; 1.
 CC InterPro; IPR004506; TrmU.
 CC Pfam; PF03054; tRNA Me trans; 1.
 CC TIGRFAMs; TIGR00420; trmU; 1.
 CC Transferrase; Methyltransferase; tRNA processing; Complete proteome.
 KW SEQUENCE 383 AA; 42709 MW; DBA2D6452CE3B78A CRC64;
 SQ
 Query Match 49.4%; Score 42; DB 1; Length 383;
 Best Local Similarity 72.7%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 SQPPEIRLTLEG 13
 Db 234 AQGEIRTVEG 244
 :|||:|||||
 RESULT 5
 ERG HUMAN
 ID ERG_HUMAN STANDARD; PRT; 462 AA.
 AC P11308;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Transcriptional regulator ERG (Transforming protein ERG).
 GN ERG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN SEQUENCE FROM N.A. (ISOFORM ERG-2).
 RP MEDLINE=87263429; PubMed=3299708;
 RA Rao V.N., Papas T.S., Shyam E., Reddy P.;
 RT "erg, a human ets-related gene on chromosome 21: alternative splicing,
 RT polyadenylation, and translation.";
 RL Science 237:635-639(1987).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM ERG-1).
 RP MEDLINE=87317608; PubMed=3476934;
 RA Reddy E.S.P., Rao V.N., Papas T.S.;
 RT "The erg gene: a human gene related to the ets oncogene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6131-6135(1987).
 RN [3]
 RN CHROMOSOMAL TRANSLOCATION.

RX MEDLINE=94356859; PubMed=8076344;
 RA Dunn T., Praissman L., Hegag N., Viola M.V.;
 RT "ERG gene is translocated in an Ewing's sarcoma cell line.";
 RL Cancer Genet. Cytogenet. 76:19-22(1994).
 RN [4]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE=94243799; PubMed=8187069;
 RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;
 RT "An RNA-binding protein gene, TIS/FUS, is fused to ERG in human
 RT myeloid leukemia with t(16;21) chromosomal translocation.";
 RL Cancer Res. 54:2865-2868(1994).
 CC -1- FUNCTION: Transcriptional regulator. May participate in
 CC transcriptional regulation through the recruitment of SETDB1
 CC histone methyltransferase and subsequent modification of local
 CC chromatin structure.
 CC -1- SUBUNIT: Interacts with SETDB1 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=ERG-2;
 CC IsoId=P11308-1; Sequence=Displayed;
 CC Name=ERG-1;
 CC IsoId=P11308-2; Sequence=VSP_001459;
 CC -1- DISEASE: Involved in a form of acute myeloid leukemia (AML)
 CC through a chromosomal translocation t(16;21)(p11;q22) that
 CC involves FUS and ERG.
 CC -1- SIMILARITY: Belongs to the ETS family.
 CC -1- SIMILARITY: Contains 1 pointed (PNT) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M17254; AAA52398.1; -;
 CC EMBL; M21535; AAA35811.1; -;
 CC PIR; A94294; TVHUEG.
 CC HSSP; Q01543; IFLI.
 CC TRANSFAC; T00265; -;
 CC TRANSFAC; T02129; -;
 CC TRANSFAC; T02130; -;
 CC Genew; HGNC:3446; ERG.
 CC MIM; 165080; -;
 CC GO; GO:0005634; C:nucleus; TAS.
 CC GO; GO:0003677; F:DNA binding; TAS.
 CC GO; GO:0004871; F:signal transducer activity; TAS.
 CC GO; GO:0008283; P:cell proliferation; TAS.
 CC GO; GO:0007275; P:development; TAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC GO; GO:0007155; P:signal transduction; TAS.
 CC InterPro; IPR000418; Ets.
 CC InterPro; IPR002341; HSF ETS.
 CC InterPro; IPR003118; SAM_PNT.
 CC Pfam; PF00178; Ets; 1.
 CC Pfam; PF02198; SAM_PNT; 1.
 CC PRINTS; PR00454; ETSDOMAIN.
 CC SMART; SM00413; ETS; 1.
 CC SMART; SM00251; SAM_PNT; 1.
 CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
 CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
 CC PROSITE; PSS0061; ETS_DOMAIN_3; 1.
 CC Transcription regulation; Nuclear protein; DNA-binding;
 KW Alternative splicing; Proto-oncogene; Chromosomal translocation.
 FT DOMAIN 122 206 POINTED.
 FT DNA_BIND 294 374 ETS-DOMAIN.
 FT VARSPPLIC 1 99 Missing (in isoform ERG-1).
 FT /FTID=VSP_001469.
 SQ SEQUENCE 462 AA; 52031 MW; B29F14B0F5C2C697 CRC64;

Query Match 49.4%; Score 42; DB 1; Length 462;
 Best Local Similarity 46.7%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WVSQPPERTLEGSC 15
 Db 63 WLSQPPARVTKMEC 77

RESULT 6

ERG_CHICK
 ID _ERG_CHICK STANDARD; PRT; 478 AA.
 AC Q30837;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcriptional regulator Erg.
 GN ERG.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archozoa; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=95329425; PubMed=7605748;
 RA Dhordain P., Dewitte F., Debliens X., Stehelin D.,
 RA Dutreque-Coquillaud M.;
 RT "Mesodermal expression of the chicken erg gene associated with
 RT precartilaginous condensation and cartilage differentiation.";
 RL Mech. Dev. 50:17-28 (1995).
 CC -1- FUNCTION: Acts as a transcriptional activator.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MESODERM- AND, TO A LESSER
 CC EXTENT, IN ECTODERM-DERIVED TISSUES.
 CC -1- SIMILARITY: Belongs to the ETS family.
 CC -1- SIMILARITY: Contains 1 pointed (PNT) domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X77159; CAA54404.1; --
 DR PIR; S60754; S60754.
 DR HSP; Q01543; IFLI.
 DR InterPro; IPR000418; Ets.
 DR InterPro; IPR002341; HSF_ETs.
 DR InterPro; IPR003118; SAM_PNT.
 DR Pfam; PF00178; Ets; 1.
 DR Pfam; PF02198; SAM_PNT; 1.
 DR PRINTS; PR00454; ETSDOMAIN.
 DR SMART; SM00413; ETS; 1.
 DR SMART; SM00251; SAM_PNT; 1.
 DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
 DR PROSITE; PS00661; ETS_DOMAIN_3; 1.
 KW Transcription regulation; Activator; Nuclear protein; DNA-binding.
 FT DOMAIN 115 199 POINTED.
 FT DNA_BIND 310 390 ETS-DOMAIN.
 FT SEQUENCE 478 AA; 53913 MW; 8DFDB243ED8623A7 CRC64;

Query Match 49.4%; Score 42; DB 1; Length 478;
 Best Local Similarity 46.7%; Pred. No. 13;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WVSQPPERTLEGSC 15
 Db 56 WLSQPPARVTKMEC 70

RESULT 7

NXSI_MICNI
 ID _NXSI_MICNI STANDARD; PRT; 60 AA.
 AC P80548;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Short neurotoxin 1 (Neurotoxin alpha).
 OS Micrurus nigrocinctus (Central American coral snake) (Gargantilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Elapidae; Serpentes; Colubroidea;
 OC Elapidae; Elapinae; Micrurus.
 OC NCBI_TaxID=8635;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=96248443; PubMed=8665942;
 RA Rosso J.-P., Vargas-Rosso O., Gutierrez J.-M., Rochat H., Bougis P.E.;
 RT "Characterization of alpha-neurotoxin and phospholipase A2 activities
 RT from Micrurus venoms. Determination of the amino acid sequence and
 RT receptor-binding ability of the major alpha-neurotoxin from Micrurus
 RT nigrocinctus nigrocinctus.";
 RL Eur. J. Biochem. 238:231-239 (1996).
 CC -1- FUNCTION: Produces peripheral paralysis by blocking neuromuscular
 CC transmission at the postsynaptic site. Binds to the nicotinic
 CC acetylcholine receptor.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -1- SIMILARITY: Belongs to the snake toxin family.
 DR PIR; S68769; S68769.
 DR HSP; P01426; INEA.
 DR InterPro; IPR003571; Snake_toxin.
 DR Pfam; PF00087; toxin; 1.
 DR ProDom; PD000206; Snake_toxin; 1.
 DR PROSITE; PS00272; SNAKE_TOXIN; 1.
 KW Toxin; Neurotoxin; Postsynaptic neurotoxin.
 FT DISULFID 3 22
 FT DISULFID 17 39 BY SIMILARITY.
 FT DISULFID 41 52 BY SIMILARITY.
 FT DISULFID 53 58 BY SIMILARITY.
 SQ SEQUENCE 60 AA; 6582 MW; 4E3580F886F0279C CRC64;

Query Match 48.8%; Score 41.5; DB 1; Length 60;
 Best Local Similarity 64.3%; Pred. No. 1.8;
 Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 3 SQPPEIRTL-EGSC 15
 Db 9 SQPPTIKTCEGQC 22

RESULT 8

YTXD_BACSU
 ID _YTXD_BACSU STANDARD; PRT; 272 AA.
 AC P39063;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein ytxd.
 GN YTXD OR BSU29730.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020526; PubMed=7934817;
 RA Grundy F.J., Waters D.A., Takova T.Y., Henkin T.M.;
 RT "Identification of genes involved in utilization of acetate and
 RT acetoin in Bacillus subtilis.";
 RL Mol. Microbiol. 10:259-271 (1993).


```

RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RL in the 200 kb rrnB-dhbA region.";
RL Microbiology 143:3431-3441(1997).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medevic V.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadajie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zunstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: May be involved in some transport function.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the motA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: LJ7309; AAC68282.1; -.
CC DR EMBL: AF008220; AAC0300.1; -.
CC DR EMBL: Z93119; CAB14951.1; -.
CC DR PIR: S39641; S39641.
CC DR Subtilist; BG10365; YtXd.
CC DR InterPro: IPR000540; Flag_MotA.
CC DR InterPro: IPR002898; MotA_ExbB.
CC DR Pfam: PF01618; MotA_ExbB; 1.
CC DR PROSITE: PS01307; MOT_A; 1.
CC DR
CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT DOMAIN 209 272 CYTOPLASMIC (POTENTIAL).
FT SEQUENCE 272 AA; 30143 MW; 376012B4BD8A7490 CRC64;
SQ
Query Match 48.2%; Score 41; DB 1; Length 272;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RL in the 200 kb rrnB-dhbA region.";
RL Microbiology 143:3431-3441(1997).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medevic V.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadajie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zunstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: May be involved in some transport function.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the motA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: LJ7309; AAC68282.1; -.
CC DR EMBL: AF008220; AAC0300.1; -.
CC DR EMBL: Z93119; CAB14951.1; -.
CC DR PIR: S39641; S39641.
CC DR Subtilist; BG10365; YtXd.
CC DR InterPro: IPR000540; Flag_MotA.
CC DR InterPro: IPR002898; MotA_ExbB.
CC DR Pfam: PF01618; MotA_ExbB; 1.
CC DR PROSITE: PS01307; MOT_A; 1.
CC DR
CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT DOMAIN 209 272 CYTOPLASMIC (POTENTIAL).
FT SEQUENCE 272 AA; 30143 MW; 376012B4BD8A7490 CRC64;
SQ
Query Match 48.2%; Score 41; DB 1; Length 272;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Best Local Similarity 42.9%; Pred. No. 11;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WVSQPPEIRTELEGS 14
Db 255 WQKQPKXQVTKKGS 268

RESULT 9
PPPS NEUCR STANDARD; PRT; 347 AA.
AC Q92250;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Farnesyl pyrophosphate synthetase (fpp synthetase) (FPP) (Farnesyl
DE pyrophosphate synthetase) [Includes: Dimethylallyltransferase
DE (EC 2.5.1.1); Geranyltransferase (EC 2.5.1.10)].
GN FPPS.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=96337905; PubMed=8753652;
RA Homann V., Mendez K., Arntz C., Ilardi V., Macino G., Morelli G.,
RA Bose G., Tudzynski B.;
RT "The isoprenoid pathway: cloning and characterization of fungal FPPS
genes";
RL Curr. Genet. 30:232-239(1996).
CC -!- FUNCTION: Catalyzes the sequential condensation of isopentenyl
CC pyrophosphate with the allylic pyrophosphates, dimethylallyl
CC pyrophosphate, and then with the resultant geranylpyrophosphate
CC to the ultimate product farnesyl pyrophosphate.
CC -!- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl
CC diphosphate = diphosphate + geranyl diphosphate.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate
CC = diphosphate + trans,trans-farnesyl diphosphate.
CC -!- PATHWAY: Isoprene biosynthesis, cholesterol biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the FPP/GGPP synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X96944; CAA65645.1; -.
CC DR PIR: S71436; S71436.
CC DR HSP: P08836; IUBX.
CC DR InterPro: IPR000092; Polyrenyl synt.
CC DR InterPro: IPR008949; Terpenoid synth.
CC DR Pfam: PF00348; polyrenyl synt; 1.
CC DR PROSITE: PS00444; POLYPRENYL SYNTHET 2; 1.
CC DR PROSITE: PS00723; POLYPRENYL SYNTHET 1; 1.
CC DR TRANSFERASE; Isoprene biosynthesis; Cholesterol biosynthesis.
KW ACT_SITE 184 184 BY SIMILARITY.
FT ACT_SITE 184 184
SQ SEQUENCE 347 AA; 40322 MW; 0D353712FB889ECE CRC64;
Query Match 48.2%; Score 41; DB 1; Length 347;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 WVSQPPEIRTELEGS 15
Db 37 WYKQSLVNTLGKGC 51

```

RESULT 10
 ID LEFT_MOUSE STANDARD; PRT; 368 AA.
 AC P57785;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Left-right determination factor B precursor (Lefty-2 protein).
 GN LEFT OR LEFTY2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98006264; PubMed=9348041;
 RA Meno C., Ito Y., Saijoh Y., Mateuda Y., Tashiro K., Kuhara S.,
 RA Hamada H.;
 RT "Two closely-related left-right asymmetrically expressed genes, lefty-
 RT 1 and lefty-2: their distinct expression domains, chromosomal linkage
 RT and direct neutralizing activity in Xenopus embryos.";
 RL Genes Cells 2:513-524(1997).
 CC -|- FUNCTION: MORPHOGEN FOR LEFT-RIGHT ASYMMETRY DETERMINATION OF
 CC ORGAN SYSTEMS IN MAMMALS.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC EXPRESSED IN THE EMERGING MESODERM. BY ES.0, EXPRESSED EXCLUSIVELY
 CC ON THE LEFT SIDE OF DEVELOPING MESODERM. WITH EXPRESSION
 CC PREDOMINANTLY IN THE LATERAL-PLATE MESODERM (LPM). WEAK EXPRESSION
 CC IN THE PROSPECTIVE FLOOR PLATE (FPF).
 CC -|- PTM: The processing of the protein may also occur at the second R-
 CC X-X-R site located at AA 132-135. Processing appears to be
 CC regulated in a cell-type specific manner.
 CC -|- SIMILARITY: Belongs to the TGF-beta family.
 DR HSP; P10600; ITGJ.
 DR MGD; MGI:1277956; Leftb.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGF-beta; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal;
 KW Multigene family.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT PROPEP 22 77 OR 135 (POTENTIAL).
 FT CHAIN 78 368 LEFT-RIGHT DETERMINATION FACTOR B.
 FT DISULFID 253 266 BY SIMILARITY.
 FT DISULFID 265 318 BY SIMILARITY.
 FT DISULFID 295 353 BY SIMILARITY.
 FT DISULFID 299 355 BY SIMILARITY.
 FT CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 368 AA; 41141 MW; 56B5ED095167A6C3 CRC64;
 Query Match 48.2%; Score 41; DB 1; Length 368;
 Best Local Similarity 47.1%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;
 QY 1 WVSQPEIRTE--GSC 15
 : : : : :
 Db 283 WILEPFGFLYECVGC 299
 RESULT 11
 ID LEFT_MOUSE STANDARD; PRT; 368 AA.
 AC Q64280;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transforming growth factor beta 4 precursor (TGF-beta 4) (Lefty
 DE protein) (Lefty-1 protein) (STRA3 protein).
 GN LEFT OR LEFTY1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98156497; PubMed=9496783;
 RA Oulad-Abdelghani M., Chazaud C., Bouillet P., Mattei M.-G., Dolle P.,
 RA Chambon P.;
 RT "STRA3/lefty, a retinoic acid-inducible novel member of the
 RT transforming growth factor-beta superfamily.";
 RL Int. J. Dev. Biol. 42:23-32(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98372436; PubMed=9708731;
 RA Meno C., Shimono A., Saijoh Y., Yashiro K., Mochida K., Ohishi S.,
 RA Noji S., Kondoh H., Hamada H.;
 RT "Lefty-1 is required for left-right determination as a regulator of
 RT lefty-2 and nodal";
 RL Cell 94:287-297(1998).
 CC -|- FUNCTION: REQUIRED FOR LEFT-RIGHT AXIS DETERMINATION AS A
 CC REGULATOR OF LEFT2 AND NODAL.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- DEVELOPMENTAL STAGE: BY ES.0, EXPRESSED EXCLUSIVELY ON THE LEFT
 CC SIDE OF DEVELOPING EMBRYOS WITH EXPRESSION PREDOMINANTLY IN THE
 CC PROSPECTIVE FLOOR PLATE (FPF). WEAK EXPRESSION IN THE LATERAL-
 CC PLATE MESODERM (LPM).
 CC -|- PTM: The processing of the protein may also occur at the second R-
 CC X-X-R site located at AA 132-135. Processing appears to be
 CC regulated in a cell-type specific manner.
 CC -|- SIMILARITY: Belongs to the TGF-beta family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 DR EMBL; D83921; BAA1212.1; -;
 DR EMBL; Z73151; CAA97497.1; -;
 DR EMBL; AJ000082; CAA03909.1; -;
 DR EMBL; AJ000083; CAA03910.1; -;
 DR PIR; S67507; S67507.
 DR HSP; P10600; ITGJ.
 DR MGD; MGI:107405; Ebaf.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGF-beta; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal.
 KW SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 76 OR 135 (POTENTIAL).
 FT CHAIN 77 368 TRANSFORMING GROWTH FACTOR BETA 4.
 FT DISULFID 253 266 BY SIMILARITY.
 FT DISULFID 265 318 BY SIMILARITY.

```
FT DISULFID 295 353 BY SIMILARITY.
FT DISULFID 299 355 BY SIMILARITY.
FT CARBOHYD 158 158 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 368 AA; 41497 MW; 821DAE663C546B5F CRC64;

Query Match 48.2%; Score 41; DB 1; Length 368;
Best Local Similarity 47.1%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 WVSQPEIRTELE-GSC 15
Db 283 WILEPPGFLTYECVGC 299

RESULT 12
ID EOMD_HUMAN STANDARD; PRT; 686 AA.
AC O95936; O8TAZ2; Q9UPM7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eomesodermin homolog.
GN EOMES OR TB2
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99337662; PubMed=10407135;
RA Kimura N., Nakashima K., Ueno M., Taga T.;
RT "A novel mammalian T-box-containing gene, Tbx2, expressed in mouse
RT developing brain.";
RL Brain Res. Dev. Brain Res. 115:183-193(1999).
RN [2]
RP SEQUENCE OF 291-455 FROM N.A.
RX MEDLINE=99107806; PubMed=9888994;
RA Yi C.-H., Terrett J.A., Li Q.-Y., Ellington K., Packham E.A.,
RA Armstrong-Buisseret L., McClure P., Slingsby T., Brook J.D.;
RT "Identification, mapping and phylogenomic analysis of four new human
RT members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19.";
RL Genomics 55:10-20(1999).
RN [3]
RP SEQUENCE OF 297-686 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SURCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: Contains 1 T-box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB011038; BA383417.1; -
EMBL; AJ010280; CA37939.1; -
EMBL; BC025363; AAH25363.1; -
HSSP; P24781; 1XBR.

DR TRANSFAC; T04395; -
DR Genew; HGNC:3372; EOMES.

DR MW; 604615; -
DR GO; GO:0003677; F:DNA binding; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR InterPro; IPR008967; P53-like.

DR InterPro; IPR001699; TF-T-box.

DR Pfam; PF00907; T-box; 1.

DR PRINTS; PR00937; TBOX.

DR SMART; SMO0425; TBOX; 1.

DR PROSITE; PS01283; TBOX_1; 1.

DR PROSITE; PS01264; TBOX_2; 1.

DR PROSITE; PS0252; TBOX_3; 1.

KW Developmental protein; Transcription regulation; DNA-binding;

KW Nuclear protein.

FT DOMAIN 117 132 POLY-ALA.

FT DOMAIN 223 229 POLY-GLY.

FT DOMAIN 246 249 POLY-ALA.

FT DNA BIND 276 456 T-BOX.

FT DOMAIN 381 384 POLY-ASN.

FT CONFLICT 337 337 C -> S (IN REF. 2).

FT CONFLICT 411 411 E -> EK (IN REF. 2).

FT CONFLICT 415 415 E -> D (IN REF. 2).

FT CONFLICT 661 661 G -> E (IN REF. 3).

FT CONFLICT 685 685 S -> T (IN REF. 3).

SQ SEQUENCE 686 AA; 72612 MW; 91A9402B1C6265FD CRC64;

Query Match 48.2%; Score 41; DB 1; Length 686;

Best Local Similarity 35.7%; Pred. No. 29;

Matches 5; Conservative 5; Mismatches 4; Indels 0;

Gaps 0;

QY 1 WVSQPEIRTELEGS 14

Db 618 WIEPPSIKSLDSN 631

RESULT 13

Y297 HUMAN

ID Y297 HUMAN STANDARD; PRT; 1411 AA.

AC Q15040; O9UEG6.

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein KIAA0297/KIAA0329.

GN KIAA0297 OR KIAA0329.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97349984; PubMed=9205841;

RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,

RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. VII.

RT The complete sequences of 100 new cDNA clones from brain which can

code for large proteins in vitro.";

RL DNA Res. 4:141-150(1997).

CC -1- SIMILARITY: TO HUMAN KIAA1017.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

RT "Multiple cDNA clones encoding nuclear proteins that bind to the tax-
 RT dependent enhancer of HTLV-1: all contain a leucine zipper structure
 RT and basic amino acid domain.";
 RL ENBO J. 9:2537-2542(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM CREB-A).
 RX MEDLINE=92087371; PubMed=1966745;
 RA Weeber G., Meyer T.E., Hoeffler J.P., Habener J.F.;
 RT "Diversification of cyclic AMP-responsive enhancer binding proteins-
 RT generated by alternative exon splicing.";
 RL Trans. Assoc. Am. Physicians 103:28-37(1990).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM CREB-B).
 RX MEDLINE=89072714; PubMed=2974179;
 RA Hoeffler J.P., Meyer T.E., Yun Y., Jameson J.L., Habener J.F.;
 RT "Cyclic AMP-responsive DNA-binding protein: structure based on a
 RT cloned placental cDNA.";
 RL Science 242:1430-1433(1988).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM CREB-B).
 RX MEDLINE=91334144; PubMed=1831258;
 RA Short M.L., Manohar C.F., Furtado M.R., Ghadge G.D., Wolinsky S.M.,
 RA Thimmapaya B., Jungmann R.A.;
 RT "Nucleotide and derived amino-acid sequences of the CRE-binding
 RT proteins from rat C6 glioma and HeLa cells.";
 RL Nucleic Acids Res. 19:4290-4290(1991).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM CREB-A).
 RX TISSUE=Eye;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=93145890; PubMed=8381074;
 RA Meyer T.E., Weeber G., Lin J., Beckmann W., Habener J.F.;
 RT "The promoter of the gene encoding 3', 5'-cyclic adenosine
 RT monophosphate (cAMP) response element binding protein contains
 RT cAMP response elements: evidence for positive autoregulation of gene
 RT transcription.";
 RL Endocrinology 132:770-780(1993).
 CC -1- FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE), A
 CC SEQUENCE PRESENT IN MANY VIRAL AND CELLULAR PROMOTERS. CREB
 CC STIMULATES TRANSCRIPTION ON BINDING TO THE CRE.
 CC -1- SUBUNIT: Binds DNA as a dimer.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Name=CREB-A;
 CC IsoId=P16220-1; Sequence=Displayed;
 CC Name=CREB-B;
 CC IsoId=P16220-2; Sequence=VSP_000596;
 CC -1- PM: Stimulated by phosphorylation.
 CC -1- SIMILARITY: Belongs to the bZIP family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S72459; AAB20597.1; -
 DR EMBL; X55545; CAA39151.1; -
 DR EMBL; M34356; ABA35717.1; -
 DR EMBL; M34356; ABA35716.1; -
 DR EMBL; M27691; AAA35715.1; -
 DR EMBL; X60003; CAA42620.1; -
 DR EMBL; BC10636; AAH10636.1; -
 DR EMBL; S53724; AAD13869.1; -
 DR PIR; A37340; A35769;
 DR PIR; B37340; B35769;
 DR PIR; S22298; S22298;
 DR TRANSFAC; T00163; -
 DR TRANSFAC; T00166; -
 DR Genew; HGNC:2345; CREB1.
 DR MIM; 123810; -
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0003712; F:transcription cofactor activity; TAS.
 DR GO; GO:0003700; F:transcription factor activity; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR GO; GO:0006351; P:transcription, DNA-dependent; TAS.
 DR InterPro; IPR001630; Leuzip_CREB.
 DR InterPro; IPR003102; PKID.
 DR Pfam; PF00170; bZIP; 1.
 DR Pfam; PF02173; PKID; 1.
 DR PRINTS; PR00041; LEUZIPPRCREB.
 DR SMART; SM00338; BRLZ; 1.
 DR PROSITE; PS00217; bZIP; 1.
 DR PROSITE; PS00036; bZIP_BASIC; 1.
 KW Transcription regulation; DNA-binding; Activator; Phosphorylation;
 KW Nuclear protein; Alternative splicing.
 FT DNA_BIND 284 305 BASIC MOTIF.
 FT DOMAIN 311 332 MISSING (in isoform CREB-B).
 FT VARSPLIC 88 101 /FTID=VSP_000596.
 FT CONFLICT 4 4 E -> D (IN REF. 5).
 FT CONFLICT 8 8 E -> D (IN REF. 5).
 FT CONFLICT 160 160 T -> A (IN REF. 5).
 FT CONFLICT 167 167 T -> A (IN REF. 5).
 FT CONFLICT 169 169 T -> A (IN REF. 5).
 FT CONFLICT 176 176 Q -> R (IN REF. 5).
 FT CONFLICT 184 184 A -> T (IN REF. 5).
 FT CONFLICT 188 188 G -> R (IN REF. 5).
 FT CONFLICT 195 195 N -> S (IN REF. 5).
 FT CONFLICT 210 210 T -> A (IN REF. 5).
 SQ SEQUENCE 341 AA; 36688 MW; D5E989AE40BF69AF CRC64;
 Query Match 47.1%; Score 40; DB 1; Length 341;
 Best Local Similarity 42.9%; Pred. No. 21;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 2 VSQPEIRTLGGSC 15
 DB 77 VIQSPQVQTVOSSC 90
 Search completed: February 26, 2004, 12:10:54
 Job time : 3.27321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 6.40584 Seconds
(without alignments)
738.822 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPEIRTLGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	89.4	115	6 Q8MJ00	Q8mj00 macaca mula
2	76	89.4	151	6 Q8MJ01	Q8mj01 macaca mula
3	76	89.4	152	4 Q95668	Q95668 homo sapien
4	76	89.4	165	4 Q95669	Q95669 homo sapien
5	76	89.4	176	4 Q95667	Q95667 homo sapien
6	76	89.4	176	6 Q95JB8	Q95JB8 macaca fasc
7	76	89.4	177	4 Q9530	Q9530 macaca mula
8	76	89.4	180	6 Q8MJ02	Q8mj02 macaca mula
9	76	89.4	190	4 Q14932	Q14932 homo sapien
10	76	89.4	201	4 Q14931	Q14931 homo sapien
11	62	72.9	192	11 Q8CG11	Q8cg11 rattus norv
12	62	72.9	192	11 Q8OWM8	Q8owm8 rattus norv
13	62	72.9	192	11 Q8CFD9	Q8cf9 rattus norv
14	49	57.6	2201	5 Q8ENN1	Q8enn1 drosophila
15	49	57.6	2529	5 Q9VWC0	Q9vwc0 drosophila
16	48	56.5	1512	2 Q93HF2	Q93hf2 streptomyce

17	48	56.5	1517	16	Q826K7	Q826k7 streptomyce
18	46.5	54.7	155	16	Q92GC3	Q92gc3 rickettsia
19	46.5	54.7	155	16	Q92CH2	Q92ch2 rickettsia
20	46	54.1	236	8	Q9TCC7	Q9tcc7 nephrolepim
21	46	54.1	397	10	Q84QL3	Q84ql3 oryza sativ
22	45	52.9	196	10	Q8H8H6	Q8h8h6 oryza sativ
23	45	52.9	353	10	Q80449	Q80449 arabidopsis
24	45	52.9	353	10	Q8LCL0	Q8lcl0 arabidopsis
25	44	51.8	650	16	Q8FWS2	Q8fws2 bruceella su
26	44	51.8	688	11	Q9JUL1	Q9jll1 mus musculu
27	44	51.8	707	11	Q8EN22	Q8bn22 mus musculu
28	43.5	51.2	804	4	Q8TER5	Q8ter5 homo sapien
29	43	50.6	305	16	Q9WX21	Q9wx21 streptomyce
30	43	50.6	456	13	Q9W6Z9	Q9w6z9 xenopus lae
31	43	50.6	474	2	Q9ADU2	Q9adu2 salmonella
32	43	50.6	485	13	Q9W700	Q9w700 xenopus lae
33	43	50.6	514	2	Q4111	Q4111 anabaena sp
34	43	50.6	514	16	Q8YMY3	Q8ywy3 anabaena sp
35	43	50.6	687	16	Q8YF69	Q8yfp69 anabaena sp
36	42	49.4	101	16	Q9XP23	Q9rx23 deinococcus
37	42	49.4	360	11	Q9D6W8	Q9d6w8 mus musculu
38	42	49.4	451	13	Q8UUU0	Q8uuu0 gallus gall
39	42	49.4	455	11	Q91XV5	Q91xv5 rattus norv
40	42	49.4	479	4	Q8IXK9	Q8ixk9 homo sapien
41	42	49.4	535	16	P73052	P73052 synecocyst
42	42	49.4	644	4	Q9NRK6	Q9nrr6 homo sapien
43	42	49.4	976	16	Q8FTA0	Q8fta0 corynebacte
44	42	49.4	1317	16	Q910F4	Q910f4 pseudomonas
45	41	48.2	91	10	Q84VG8	Q84vg8 arabidopsis

ALIGNMENTS

RESULT 1

Q8MJ00 PRELIMINARY; PRT; 115 AA.

AC Q8MJ00; 01-OCT-2002 (T-EMBLrel. 22, Created)

DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (T-EMBLrel. 25, Last annotation update)

DE NKP30S.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Macaca.

OC Cercopitheidae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RA LaBonte M.L.; Miller J.; Letvin N.L.;

RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification of NKP46SD and NKP30S."

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY035217; AXK63119.1; -

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR Pfam; PF00047; IG; 1.

DR SMART; SM00409; IG; 1.

DR PROSITE; PS50835; IG LIKE; 1.

SQ SEQUENCE 115 AA; 12871 MW; 1D34CC0E8986DE9F CRC64;

Query Match 89.4%; Score 76; DB 6; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRTLGGS 14
|||

Db 20 WVSQPPPEIRTLGGS 33
|||

RESULT 2

Q8MJ01 PRELIMINARY; PRT; 151 AA.

AC Q9MJ01;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE NKp30v1ED
 OS Macaca mulatta (Rhesus macaque).
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LaBonte M.L., Miller J., Letvin N.L.;
 RT "Molecular cloning of rhesus monkey NKp46 and identification
 RT of NKp46SD and NKp30S.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY035216; AAK63118.1; --
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; ig; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;
 Query Match 89.4%; Score 76; DB 6; Length 151;
 Best Local Similarity 100.0%; Pred No. 2e-05; 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0;
 QY 1 WVSQPPPIRLTLEGS 14
 DB |||||
 20 WVSQPPPIRLTLEGS 33
 RESULT 3
 ID O95668 PRELIMINARY; PRT; 152 AA.
 AC O95668;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7d.
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bouguetel L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NFKAPPAB
 RT family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93208881; PubMed=7916655;

RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pemica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Rethmüller G., Weiss E.H.;
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
 RT Level of TNF-beta Production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [7]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91139175; PubMed=1671667;
 RA Abraham I.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [8]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94382679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
 RT Nuclear RNA Helicase of the D-E-A-D Family.";
 RL Genomics 26:210-218(1995).
 RN [10]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [11]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [12]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98149985; PubMed=9480751;
 RA Shima T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT LKB1 and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 DR EMBL; Y14768; CAA75067.1; --

DR GO: GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR GO: GO:0006954; P:inflammatory response; NAS.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR007110; IG-like.
 DR Pfam: PF00047; IG; 1.
 DR SMART: SM00403; IG; 1.
 DR PROSITE: PS00835; IG-LIKE; 1.
 SQ SEQUENCE 152 AA; 16393 MW; 42718746451F9ADC CRC64;

Query Match 89.4%; Score 76; DB 4; Length 152;
 Best Local Similarity 100.0%; Pred. No. 2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGS 14
 DB 20 WVSQPPERTLEGS 33
 |||||
 |||||

RESULT 4

ID 095669 PRELIMINARY; PRT; 165 AA.

AC 095669;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7F.
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bouguetel L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.,
 RT "Dense Alu clustering and a potential new member of the NFKappaB
 family within a 90 Kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brien-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a
 heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [6]

RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Rietmuller G., Weiss E.H.;
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
 Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
 Correlates with A Variant Amino Acid in Position 26 and a Reduced
 Level of TNF-beta Production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95324911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
 RT Nuclear RNA Helicase of the D-E-A-D Family.";
 RL Genomics 26:210-218(1995).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IkBL and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 DR EMBL: Y14768; CAA75068.1;
 DR GO: GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR GO: GO:0006954; P:inflammatory response; NAS.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR007110; IG-like.
 DR Pfam: PF00047; IG; 1.
 DR SMART: SM00403; IG; 1.
 DR PROSITE: PS00835; IG-LIKE; 1.
 SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3B1E7E CRC64;

Query Match 89.4%; Score 76; DB 4; Length 165;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1: WVSQPPEIRTELEGS 14
 Db |||||

20 WVSQPPEIRTELEGS 33

RESULT 5

ID O95667 PRELIMINARY; PRT; 176 AA.
 AC O95667;
 DT 01-MAY-1999 (T-EMBLrel. 10, Created)
 DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE 1C7E.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bouguieret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.,
 RT "Dense Alu clustering and a potential new member of the NFKAPPAB
 family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage
 molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region";
 RL Immunogenetics 42:315-322(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Graco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a
 heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Rietmueller G., Weiss E.H.;
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced
 RT Level of TNF-beta Production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";

RL Immunogenetics 33:50-53(1991).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [9]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9534911; PubMed=7601445;
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
 RA Van Zevenen A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative
 RT Nuclear RNA Helicase of the D-E-A-D Family.";
 RL Genomics 26:210-218(1995).
 RN [10]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [11]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [12]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98149985; PubMed=9480751;
 RA Shilina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IKB and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 DR EMBL: Y14768; CAA75066.1;
 DR GO: GO:0003793; P:defense/immunity protein activity; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR GO: GO:0006954; P:inflammatory response; NAS.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR007110; IG-like.
 DR Pfam: PF00047; IG; 1.
 DR SMART: SM00409; IG; 1.
 DR PROSITE: PS50835; IG LIKE; 1.
 SQ SEQUENCE 176 AA; 18749 MW; 162BBB775DA2BCD35 CRC64;
 Query Match 89.4%; Score 76; DB 4; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVSQPPEIRTELEGS 14
 Db |||||

RESULT 6

Q95JB8 PRELIMINARY; PRT; 176 AA.
 ID O95JB8
 AC O95JB8;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE NFKB30 (NFKB30v1).
 GN NCR.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541, 9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.fascicularis; TISSUE=Lymphoid;
 RA Rizzi M., Biassoni R.;
 RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in
 RT Macaca fascicularis lymphoid cells.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mulatta;
 RA LaBonte M.L., Miller J., Letvin N.L.;
 RT "Molecular cloning of rhesus monkey NKP46 and identification
 RT of NKP46SD and NKP30S.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ278389; CAC41081.1; -;
 DR EMBL; AY035215; AAK63117.1; -;
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;
 Query Match 89.4%; Score 76; DB 6; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVSQPEPTLRLGSS 14
 Db 20 WVSQPEPTLRLGSS 33
 RESULT 7
 O14930
 ID O14930 PRELIMINARY; PRT; 177 AA.
 AC O14930;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 1C7 precursor (1C7 protein).
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";
 RL Genomics 31:215-222(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=9327029; PubMed=8499947;
 RA Irls F., Bouguetere L., Prieur S., Caterina D., Perrot V.,
 RA Jurka J., Rodriguez-Tone P., Clavierie J., Cohen D., Dausset J.;
 RT "Dense Alu clustering and a potential new member of the NF kappa B
 RT family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts.";
 RL Transplantation 61:1387-1392(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierrina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotixin beta, a novel member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
 RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotixin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91086846; PubMed=1670638;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Riethmuller G., Weiss E.H.;
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced
 RT level of TNF-beta production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [9]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=91139175; PubMed=1671667;
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RT "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [10]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [11]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95324911; PubMed=7601445;
 RA Feilman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vainan M.,
 RA Van Zeveren A., Coppieters W., Van de Waghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family.";
 RL Genomics 26:210-218(1995).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20132445; PubMed=10668961;
 RA Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:155-160(1997).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Feilerhoff B., Maier S., Martinozzi S., Weidie U.,
 RA Weiss E.H.;

RT "Complex expression pattern of the TNF region gene LST1 through
 RL differential regulation, initiation, and alternative splicing."
 RL Genomics 45:591-600(1997).
 RN [14]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9814985; PubMed=9480751;
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Tsuchi S., Sugawara C.,
 RA Oka A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikenura T.,
 RA Kimura M., Inoko H.;
 RA "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IRL and MICA genes at the centromeric end of the HLA class I
 RT region."
 RL Genomics 47:372-382(1998).
 RN [15]
 RN SEQUENCE FROM N.A.
 RP HIRAKAWA M., YAMAGUCHI H., IMAI K., SHIMADA J.;
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL [16]
 RN SEQUENCE FROM N.A.
 RA Shiina S., Tamiya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031136; AB86578.1; -;
 DR EMBL; Y14768; CAAV5064.1; -;
 DR EMBL; AP00505; BAB63393.1; -;
 DR Genew; HGNC:14189; LST1.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0003793; P:defense/immunity protein activity; TAS.
 DR GO; GO:0006968; P:cellular defense response; TAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS08335; IG_LIKE; 1.
 KW SIGNAL.
 FT CHAIN 1 61 POTENTIAL.
 FT CHAIN 62 177 IC7.
 SQ SEQUENCE 177 AA; 19237 MW; DD5EC96F0AB2DCE6 CRC64;
 Query Match 89.4%; Score 76; DB 6; Length 177;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVSQPPPIRTLEGS 14
 Db 20 WVSQPPPIRTLEGS 33
 RESULT 8
 ID QBMJ02 PRELIMINARY; PRT; 180 AA.
 AC QBMJ02;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Nkp30.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA LaBonte M.L., Miller J., Letvin N.L.;
 RT "Molecular cloning of rhesus monkey Nkp46 and identification
 RT of Nkp46SD and Nkp30S."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY035214; AAK63116.1; -;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.

DR PROSITE; PS08335; IG_LIKE; 1.
 SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;
 Query Match 89.4%; Score 76; DB 6; Length 180;
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WVSQPPPIRTLEGS 14
 Db 20 WVSQPPPIRTLEGS 33
 RESULT 9
 ID O14932 PRELIMINARY; PRT; 190 AA.
 AC O14932;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE IC7 precursor (Natural killer cell receptor).
 GN IC7 OR NKP30.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96422187; PubMed=8824804;
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."
 RL Genomics 31:215-222(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
 RT "Jense Alu clustering and a potential new member of the NF kappa B
 family within a 90 kilobase HLA class III segment."
 RL Nat. Genet. 3:137-145(1993).
 RN [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96215741; PubMed=8629302;
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
 RA Wallace A.F., Russell M.E.;
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage
 RT molecule expressed in transplanted human hearts."
 RL Transplantation 61:1387-1392(1996).
 RN [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
 RA Weiss E.H.;
 RT "Cloning and genomic characterization of LST1: a new gene in the human
 RT TNF region."
 RL Immunogenetics 42:315-322(1995).
 RN [6]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a
 RT heteromeric complex with lymphotoxin on the cell surface."
 RL Cell 72:847-856(1993).
 RN [7]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86016093; PubMed=2995927;
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,

RA Pennica D., Goeddel D.V., Gray P.W.;
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,
 RT homology and chromosomal localization.";
 RL Nucleic Acids Res. 13:6361-6373(1985).
 RN [8]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=91086846; PubMed=1670639;
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
 RA Rietmueller G., Weiss E.H.;
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
 RT NcoI polymorphism in the first intron of the human TNF-beta gene
 RT correlates with a variant amino acid in position 26 and a reduced
 RT level of TNF-beta production.";
 RL J. Exp. Med. 173:209-219(1991).
 RN [9]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=91139175; PubMed=1671667;
 RX Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
 RA "Haplotypic polymorphisms of the TNFB gene.";
 RL Immunogenetics 33:50-53(1991).
 RN [10]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=94362679; PubMed=8081366;
 RA Albertella M.R., Campbell D.R.;
 RT "Characterization of a novel gene in the human major
 RT histocompatibility complex that encodes a potential new member of the
 RT I kappa B family of proteins.";
 RL Hum. Mol. Genet. 3:793-799(1994).
 RN [11]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=95324911; PubMed=7601445;
 RX Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
 RA Strominger J., Spies T.;
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative
 RT nuclear RNA helicase of the DEAD family.";
 RL Genomics 26:210-218(1995).
 RN [12]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=20132445; PubMed=10668961;
 RX Neville M.J., Campbell R.D.;
 RT "Alternative splicing of the LST-1 gene located in the major
 RT histocompatibility complex on human chromosome 6.";
 RL DNA Seq. 8:153-160(1997).
 RN [13]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=98035883; PubMed=9367684;
 RX de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidie U.,
 RA Weiss E.H.;
 RT "Complex expression pattern of the TNF region gene LST1 through
 RT differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [14]
 RA SEQUENCE FROM N.A.
 RP MEDLINE=98149985; PubMed=9480751;
 RX Shiina T., Taniya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
 RA Kimura M., Inoko H.;
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the
 RT IkBL and MICA genes at the centromeric end of the HLA class I
 RT region.";
 RL Genomics 47:372-382(1998).
 RN [15]
 RA SEQUENCE FROM N.A.
 RP TISSUE=Lymphoid;
 RX Blassoni R., Pessino A., Malaspina A.;
 RA "NK-1 activating NK receptor.";
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RA SEQUENCE FROM N.A.
 RP TISSUE=Peripheral blood;
 RX Sato M., Yabe T., Ohashi J., Tsuchiya N., Hanaoka K., Tokunaga K.,

RA Juji T.;
 RT "Identification of two novel single nucleotide polymorphisms in the
 RT Nkp30 gene in human natural killer cells.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [17]
 RA EMBL; AF031138; AAB86580.1; -;
 DR EMBL; Y14768; CAA75065.1; -;
 DR EMBL; AJ223153; CAB54004.1; -;
 DR EMBL; AB055881; BAB78472.1; -;
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR CO; GO:0006954; P:inflammatory response; NAS.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS0835; IG_LIKE; 1.
 KW Signal; Receptor.
 FT SIGNAL 1 61 POTENTIAL.
 FT CHAIN 62 190 1C7.
 SQ SEQUENCE 190 AA; 20640 MW; 1FAC919E20A6B18A CRC64;
 Query Match 89.4%; Score 76; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 2.6e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 WVSQPPPIRTLEGS 14
 Db 20 WVSQPPPIRTLEGS 33
 RESULT 10
 O14931
 ID O14931 PRELIMINARY; PRT; 201 AA.
 AC O14931;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE 1C7 precursor (NCR3 protein).
 GN 1C7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=96422187; PubMed=8924804;
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";
 RL Genomics 31:215-222(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Olin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;
 RT "Sequence of the human major histocompatibility complex class III
 RT region.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93272029; PubMed=8499947;
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,
 RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;
 RT "Pense Alu clustering and a potential new member of the NF kappa B
 RT family within a 90 kilobase HLA class III segment.";
 RL Nat. Genet. 3:137-145(1993).
 RN [5]
 RP SEQUENCE FROM N.A.

RA MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,
RA Wallace A.F., Russell M.E.;
RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts.";
RL Transplantation 61:1387-1392 (1996).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwiertzina H.,
RA Weiss E.H.;
RT "Cloning and genomic characterization of LST1: a new gene in the human
RT TNF region";
RL Immunogenetics 42:315-322 (1995).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=9320881; PubMed=7916655;
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;
RT "Lymphotoxin beta: a novel member of the TNF family that forms a
RT heteromeric complex with lymphotoxin on the cell surface.";
RL Cell 72:847-856 (1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=86016093; PubMed=2995927;
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,
RA Pennica D., Goeddel D.V., Gray P.W.;
RT "Human lymphotoxin and tumor necrosis factor genes: structure,
RT homology and chromosomal localization.";
RL Nucleic Acids Res. 13:6361-6373 (1985).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=91086846; PubMed=1670638;
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,
RA Riethmuller G., Weiss E.H.;
RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an
RT NcoI polymorphism in the first intron of the human TNF-beta gene
RT correlates with a variant amino acid in position 26 and a reduced
RT level of TNF-beta production.";
RL J. Exp. Med. 173:209-219 (1991).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671657;
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;
RT "Haplotypic polymorphisms of the TNFB gene.";
RL Immunogenetics 33:50-53 (1991).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94362679; PubMed=8081366;
RA Albertella M.R., Campbell D.R.;
RT "Characterization of a novel gene in the human major
RT histocompatibility complex that encodes a potential new member of the
RT I kappa B family of proteins.";
RL Hum. Mol. Genet. 3:793-799 (1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=95324911; PubMed=7601445;
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,
RA Van Zeven A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,
RA Strominger J., Spies T.;
RT "The B2M gene in the MHC encodes an evolutionarily conserved putative
RT nuclear RNA helicase of the DEAD family.";
RL Genomics 26:210-218 (1995).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=20132445; PubMed=10668961;
RA Neville M.J., Campbell R.D.;
RT "Alternative splicing of the LST-1 gene located in the major
RT histocompatibility complex on human chromosome 6.";
RL DNA Seq. 8:155-160 (1997).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035883; PubMed=9367684;

RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,
RA Weiss E.H.;
RT "Complex expression pattern of the TNF region gene LST1 through
RT differential regulation, initiation, and alternative splicing.";
RL Genomics 45:591-600 (1997).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=98149985; PubMed=9480751;
RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,
RA Kimura M., Inoko H.;
RT "Nucleotide sequencing analysis of the 145-kilobase segment around the
RT IkBL and MICA genes at the centromeric end of the HLA class I
RT region.";
RL Genomics 47:372-382 (1998).
RN [16]
RP SEQUENCE FROM N.A.
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWhann P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.D., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fhney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [17]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139175; PubMed=1671657;
RA Strausberg R.;
RT Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
RL ENBL; AF031137; AB84579.1; -;
DR ENBL; AF129756; AAD18088.1; -;
DR ENBL; Y14768; CAAY5063.1; -;
DR ENBL; BC052582; AAH52582.1; -;
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0006954; P:inflammatory response; NAS.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 62 201 107 POTENTIAL.
SQ SEQUENCE 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;

Query Match 89.4%; Score 76; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTELEGS 14

Db 20 WVSQPPEIRTELEGS 33

RESULT 11
Q8CG11

ID Q8CG11 PRELIMINARY; PRT; 192 AA.
 AC Q8CG11; 2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE NK receptor 1c7 precursor.
 GN 1c7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BN;
 RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
 RA Butcher G.W.;
 RT "Molecular characterization of the novel rat NK receptor 1c7."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ430419; CAD23067.2;
 DR EMBL; AJ430420; CAD23067.2; JOINED.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS0835; IG-LIKE; 1.
 KW SIGNAL; Receptor.
 FT SIGNAL 1 19 POTENTIAL.
 SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3AFBE6DC0 CRC64;
 Query Match 72.9%; Score 62; DB 11; Length 192;
 Best Local Similarity 78.6%; Pred. No. 0.0078;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVSQPEIRTLGSG 14
 Db |||||
 20 WVSQPEIRTLGSGT 33
 RESULT 12
 ID Q8OWM8 PRELIMINARY; PRT; 192 AA.
 AC Q8OWM8;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE NKp30.
 GN NKp30.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEW;
 RX MEDLINE=22168131; PubMed=12180816;
 RA Hsieh C.L., Ohara H., Ogura Y., Martinez O.M., Krams S.M.;
 RT "NK cells and translocation."
 RL Transpl. Immunol. 9:111-114(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEW;
 RA Hsieh C.L., Ohara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
 RA Martinez O.M., Krams S.M.;
 RT "Identification, Cloning, and Characterization of a Novel Rat NK
 Receptor, rNKp30: a Molecule Expressed in Liver Allografts."
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY273824; AAP13457.1;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR SMART; SM00409; IG; 1 IG-LIKE.
 DR PROSITE; PS0835; IG-LIKE; 1.
 SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

Query Match 72.9%; Score 62; DB 11; Length 192;
 Best Local Similarity 78.6%; Pred. No. 0.0078;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVSQPEIRTLGSG 14
 Db |||||
 20 WVSQPEIRTLGSGT 33
 RESULT 13
 ID Q8CFD9 PRELIMINARY; PRT; 192 AA.
 AC Q8CFD9;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE 1c7 protein precursor.
 GN 1c7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PVG; TISSUE=NK cell;
 RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
 RA Butcher G.W.;
 RT "Molecular characterization of the novel rat NK receptor 1c7."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ430418; CAD23066.1;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR Pfam; PF00047; IG; 1.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS0835; IG-LIKE; 1.
 KW SIGNAL.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 192 IC7 PROTEIN.
 SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;
 Query Match 72.9%; Score 62; DB 11; Length 192;
 Best Local Similarity 78.6%; Pred. No. 0.0078;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 WVSQPEIRTLGSG 14
 Db |||||
 20 WVSQPEIRTLGSGT 33
 RESULT 14
 ID Q86NN1 PRELIMINARY; PRT; 2201 AA.
 AC Q86NN1;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE GH07949p (Fragment).
 GN CG11936.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleib J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;

```

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT004503; AAC42667.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001025; BAH.
DR SMART; SM00439; BAH; 1.
FT NONTER
SQ SEQUENCE 2201 AA; 237100 MW; 10303133A8B4864B CRC64;

Query Match 57.6%; Score 49; DB 5; Length 2201;
Best Local Similarity 81.8%; Pred. No. 20;
Matches 9; Conservative 1; Mismatches 0; Gaps 0;

QY 4 QPPEIRTLGDS 14
DB 340 QPPEIRTLGDS 350

RESULT 15
Q9VWCO PRELIMINARY; PRT; 2529 AA.
AC Q9VWCO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG32529 protein.
GN CG32529 OR CG11936 OR CG55619.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Makos G.L.G.,
RA Ballow R.M., Basu A., Baxendale J., Andrews-Frankoch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jallali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Ciamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003513; AAP49026.2; -
DR FlyBase; FBgn0052529; CG32529.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001025; BAH.
DR SMART; SM00439; BAH; 1.
SQ SEQUENCE 2529 AA; 270575 MW; C93B1726EE7547D4 CRC64;

Query Match 57.6%; Score 49; DB 5; Length 2529;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QPPEIRTLGDS 14
DB 668 QPPEIRTLGDS 678

```

Search completed: February 26, 2004, 12:13:50
Job time : 8.40584 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 9.31034 Seconds
(without alignments)
455.215 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPEIRTELESC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	4	Aae02774 Human Nkp
2	76	89.4	120	4	Aae02771 Human Nkp
3	76	89.4	135	5	Aae19109 Human Nkp
4	76	89.4	177	2	Aay06402 Human B-C
5	76	89.4	190	2	Aay06401 Human B-C
6	76	89.4	190	4	Aae02769 Human Nkp
7	76	89.4	201	2	Aay06403 Human B-C
8	76	89.4	369	5	Aae19110 Human Nkp
9	49	57.6	1289	4	Abb70840 Drosophil
10	45	52.9	164	4	Aau27977 Human con
11	45	52.9	301	3	Aag43485 Arabidops
12	45	52.9	353	3	Aag43484 Arabidops
13	45	52.9	353	5	Abb92033 Herbicida
14	45	52.9	380	3	Aag43483 Arabidops
15	44	51.8	185	4	Abg23939 Novel hum
16	44	51.8	688	3	Aay57393 Mouse ner
17	43	50.6	74	4	Aam88325 Human imm
18	43	50.6	276	4	Abg07919 Novel hum
19	43	50.6	514	5	Abj10466 XisP reco
20	42	49.4	63	5	Abp02197 Human ORF
21	42	49.4	134	3	Aag01336 Human sec
22	42	49.4	383	6	Abu39323 Protein e
23	42	49.4	451	2	Aay01520 Chicken C
24	42	49.4	459	6	Abp97695 Amino aci
25	42	49.4	462	5	Aau75313 Human Erg

ALIGNMENTS

RESULT 1

Aae02774

ID AAE02774 standard; peptide; 15 AA.

XX AAE02774;

XX 06-AUG-2001 (first entry)

XX Human Nkp30 receptor immunogenic peptide for antiserum production.

XX Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;

KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

KW therapy.

XX Homo sapiens.

XX WO200136630-A2.

XX 25-MAY-2001.

XX 15-NOV-2000; 2000WO-EP011697.

XX 15-NOV-1999; 99CA-02288307.

XX 15-NOV-1999; 99US-00440514.

XX (INNA-) INNATE PHARMA SAS.

XX (OYGE-) UNIV GENOVA.

XX Moretta A, Bottino C, Biassoni R;

XX WPI; 2001-329221/34.

XX Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.

XX Claim 1; Page 33; 83pp; English.

XX The invention relates to human Nkp30 receptor and its corresponding cDNA

XX molecule which is involved in natural cytotoxicity mediated by natural

XX killer (NK) cells and antibodies that identify the same. Nkp30 receptor

XX is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively

XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are

XX useful for detecting and/or quantifying the presence of NK cells in a

XX biological sample. The invention also provide kits for detecting and/or

XX quantifying the presence of NK cells, for the selective removal of NK

XX cells from a biological sample, for the positive and selective

Abu03476 Angiogene
Abu5656 Lung canc
Adb75275 Prostate
Aay01521 Chicken c
Aau33622 Pseudomon
Abu15571 Protein e
Abp42291 Human ova
Abp33534 Human ORF
Aab33534 Human ORF
Aau80378 Human DOP
Abg02616 Novel hum
Aay49415 Human MTB
Aay59954 Human end
Aay03852 Murine le
Aay49414 Human T-b
Abm72545 Staphyloc
Abm67581 Drosophil
Aay57394 Human ner
Abg21153 Novel hum
Aao29646 Tribolium
Abb40612 Peptide #

26 42 49.4 462 6 ABU03476
27 42 49.4 462 6 ABU56566
28 42 49.4 462 7 ADB75275
29 42 49.4 478 2 AAY01521
30 42 49.4 1317 4 AAU33622
31 42 49.4 1317 6 ABU15571
32 41.5 48.8 116 5 ABP42291
33 41 48.2 87 5 ABP33534
34 41 48.2 137 5 AAU80378
35 41 48.2 149 4 ABG02616
36 41 48.2 249 3 AAY49415
37 41 48.2 289 2 AAY59954
38 41 48.2 368 2 AAY03852
39 41 48.2 517 3 AAY49414
40 41 48.2 520 6 ABM72545
41 41 48.2 630 4 ABM67581
42 41 48.2 686 3 AAY57394
43 40.5 47.6 498 4 ABG21153
44 40.5 47.6 756 6 AAO29646
45 40 47.1 62 4 ABB40612

CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft
 CC versus tumour (GVT) and especially graft versus leukaemia (GvL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. Nkp30 antibodies are useful for
 CC identifying Nkp30 natural ligands and allow assessment of the level of
 CC surface Nkp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence Nkp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is human Nkp30 receptor immunogenic peptide fragment
 CC which is used for Nkp30 polyclonal antiserum production
 XX
 XX Sequence 15 AA;

Query Match 100.0%; Score 85; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 8.3e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGSC 15
 |||||
 Db 1 WVSQPPEIRTLGSC 15

RESULT 2
 AAE02771
 ID AAE02771 standard; protein; 120 AA.
 XX
 AC AAE02771;
 XX
 DT 06-AUG-2001 (first entry)
 XX
 DE Human Nkp30 receptor extracellular region sequence.
 XX
 KW Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
 KW therapy; extracellular region.
 XX
 OS Homo sapiens.
 XX
 PN WO200136630-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 15-NOV-2000; 2000WO-EP011697.
 XX
 PR 15-NOV-1999; 99CA-02288307.
 PR 15-NOV-1999; 99US-00440514.
 XX
 PA (INNA-) INNATE PHARMA SAS.
 PA (UYGE-) UNIV GENOVA.
 XX
 PI Moretta A, Bottino C, Biassoni R;
 XX
 DR WPI; 2001-329221/34.
 XX
 PT Novel compound, useful for detection and/or quantifying the presence of
 PT NK cells, comprises the amino acid sequences of the Nkp30 molecule.
 XX
 PS Claim 1; Fig 7B; 83pp; English.
 XX
 CC The invention relates to human Nkp30 receptor and its corresponding cDNA
 CC molecule which is involved in natural cytotoxicity mediated by natural
 CC killer (NK) cells and antibodies that identify the same. Nkp30 receptor
 CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively
 CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are
 CC useful for detecting and/or quantifying the presence of NK cells in a
 CC biological sample. The invention also provide kits for detecting and/or
 CC quantifying the presence of NK cells, for the selective removal of NK
 CC cells from a biological sample, for the positive and selective
 CC purification of NK cells from a biological sample and for the in vitro

CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft
 CC versus tumour (GVT) and especially graft versus leukaemia (GvL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. Nkp30 antibodies are useful for
 CC identifying Nkp30 natural ligands and allow assessment of the level of
 CC surface Nkp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence Nkp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is the extracellular region of human Nkp30 receptor
 XX
 XX Sequence 120 AA;

Query Match 89.4%; Score 76; DB 4; Length 120;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGSC 14
 |||||
 Db 2 WVSQPPEIRTLGSC 15

RESULT 3
 AAE19109
 ID AAE19109 standard; protein; 135 AA.
 XX
 AC AAE19109;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Human Nkp30 protein.
 XX
 KW Human; natural killer cell activating protein; Nkp46; therapy; virucide;
 KW viral infection; natural killer cell; NK; Nkp44; imaging agent; cancer;
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; Nkp30.
 XX
 OS Homo sapiens.
 XX
 PN WO200208287-A2.
 XX
 PD 31-JAN-2002.
 XX
 PF 19-JUL-2001; 2001WO-IL000664.
 XX
 PR 20-JUL-2000; 2000IL-00137419.
 XX
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA (UYNE) UNIV BEN-GURION NEGEV.
 XX
 PI Mandelboim O, Porgador A;
 XX
 DR WPI; 2002-195870/25.
 DR N-PSDB; AAD30466.
 XX
 PT New targeting complex capable of targeting an active substance to a
 PT target cell, comprising a target recognition segment and an active
 PT segment, useful for treating pathologies associated with viral infections
 PT or cancer.
 XX
 PS Example 1; Page 108; 113pp; English.
 XX
 CC The invention relates to compositions and methods for the treatment and
 CC detection of a variety of viral infections, by using complex agents
 CC comprising the natural killer (NK) cells activating proteins, Nkp46 and
 CC Nkp44 and functional fragments thereof, linked to therapeutic or imaging
 CC agents. The complex is useful for treating pathologies associated with
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECVV, MVM or herpes virus)
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
 CC the imaging and monitoring of cancer. The complex may also be used to
 CC detect the presence of abnormal cells in a sample. The antibodies can be

CC used to qualitatively or quantitatively detect the ligand for the
CC complex. The present sequence is human NRP30 protein
XX
SQ Sequence 135 AA;

Query Match 89.4%; Score 76; DB 5; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPEIRTELEGS 14
Db 20 WVSQPEIRTELEGS 33

RESULT 4
AAY06402
ID AAY06402 standard; protein; 177 AA.

XX AC AAY06402;
XX DT 20-SEP-1999 (first entry)
XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
XX KW signal transduction; immunomodulator; antiinflammatory;
XX KW autoimmune disease; inflammation; gene therapy; diagnosis.
XX OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..12
Protein	/note= "leader peptide"
Modified-site	42
Modified-site	/note= "N-glycosylated"
Modified-site	68
Modified-site	/note= "N-glycosylated"
Modified-site	121
Modified-site	/note= "N-glycosylated"
Domain	139..162
Peptide	/note= "transmembrane domain"
Peptide	166..177
Peptide	/note= "alternatively spliced C-terminal end"

WO9923867-A2.
20-MAY-1999.
05-NOV-1998; 98WO-US023826.
07-NOV-1997; 97US-0064761P.
(BIOJ) BIOGEN INC.
Browning J;
WPI; 1999-418423/35.
N-PSDB; AAX59348.
Novel B-cell myelin oligodendrocyte glycoproteins.
Claim 2; Page 43; 43pp; English.

This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of

CC producing BMOG using these transformed host cells are also provided. BMOG CC polypeptides can be used for modulating the immune system of a subject or CC to inhibit signal transduction in a cell expressing BMOG by contacting it CC with a soluble BMOG protein. The nucleic acid can be used for gene CC therapy. The protein can also be used to target a toxin, imaging agent or CC radionuclide to a cell expressing BMOG. (All claimed)

XX Sequence 177 AA;

Query Match 89.4%; Score 76; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPEIRTELEGS 14
Db 20 WVSQPEIRTELEGS 33

RESULT 5
AAY06401
ID AAY06401 standard; protein; 190 AA.

XX AC AAY06401;
XX DT 20-SEP-1999 (first entry)
XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
XX KW signal transduction; immunomodulator; antiinflammatory;
XX KW autoimmune disease; inflammation; gene therapy; diagnosis.
XX OS Homo sapiens.

Key	Location/Qualifiers
Peptide	1..12
Protein	/note= "leader peptide"
Modified-site	13..190
Modified-site	/note= "mature protein"
Modified-site	42
Modified-site	/note= "N-glycosylated"
Modified-site	68
Modified-site	/note= "N-glycosylated"
Modified-site	121
Modified-site	/note= "N-glycosylated"
Domain	139..162
Peptide	/note= "transmembrane domain"
Peptide	166..190
Peptide	/note= "alternatively spliced C-terminal end"

WO9923867-A2.
20-MAY-1999.
05-NOV-1998; 98WO-US023826.
07-NOV-1997; 97US-0064761P.
(BIOJ) BIOGEN INC.
Browning J;
WPI; 1999-418423/35.
N-PSDB; AAX59347.
Novel B-cell myelin oligodendrocyte glycoproteins.
Claim 2; Page 42; 43pp; English.

This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph

CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)

XX SQ Sequence 190 AA;

Query Match 89.4%; Score 76; DB 2; Length 190;
 Best Local Similarity 100.0%; Pred. No. 0.00033;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPPEIRTLGSS 14
 |||||
 Db 20 WVSQPPPEIRTLGSS 33

RESULT 6

AAE02769
 ID AAE02769 standard; protein; 190 AA.

XX AC AAE02769;

XX DT 06-AUG-2001 (first entry)

XX DE Human NKp30 receptor.

XX KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;
 KW therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..18 /label= Signal_peptide

FT Protein 19..190 /label= Mature_NKp30_receptor_protein

FT Region 19..138 /label= Extracellular_region

FT FT /label= Forms an immunoglobulin (Ig) V-like domain"

FT Modified-site 42 /note= "N-glycosylation site"

FT Modified-site 121 /note= "N-glycosylation site"

FT Region 139..157 /label= Transmembrane_region

FT Region 158..190 /label= Intracellular_region

XX WO200136630-A2.

XX PD 25-MAY-2001.

XX PF 15-NOV-2000; 2000WO-EP011697.

XX PR 15-NOV-1999; 99CA-02288307.

XX PR 15-NOV-1999; 99US-00440514.

XX PA (INNA-) INNATE PHARMA SAS.

XX PA (UYGE-) UNIV GENOVA.

XX PI Moretta A, Bottino C, Biassoni R;

XX DR WPI; 2001-329221/34.

XX DR N-PSDB; AAD06564.

PT Novel compound, useful for detection and/or quantifying the presence of
 PT NK cells, comprises the amino acid sequences of the NKp30 molecule.

XX Claim 1; Fig 7B; 83pp; English.

XX CC The invention relates to human NKp30 receptor and its corresponding cDNA
 CC molecule which is involved in natural cytotoxicity mediated by natural
 CC killer (NK) cells and antibodies that identify the same. NKp30 receptor
 CC is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively
 CC expressed on the surface of human mature NK cells. NKp30 and its cDNA are
 CC useful for detecting and/or quantifying the presence of NK cells in a
 CC biological sample. The invention also provide kits for detecting and/or
 CC quantifying the presence of NK cells, for the selective removal of NK
 CC cells from a biological sample, for the positive and selective
 CC purification of NK cells from a biological sample and for the in vitro
 CC stimulation of NK cell cytotoxicity. The invention further provides a
 CC pharmaceutical composition which is used as a drug for grafting
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft
 CC versus tumour (GVT) and especially graft versus leukaemia (GVL), and for
 CC the prevention, palliation and/or therapy of solid or liquid tumours,
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
 CC microorganism, notably viral infection. NKp30 antibodies are useful for
 CC identifying NKp30 natural ligands and allow assessment of the level of
 CC surface NKp30 ligand expressed on an NK-susceptible target cell and the
 CC comparison of this level to the standard physiological one. Hence NKp30
 CC antibodies are useful in the diagnosis of tumours or of infection. The
 CC present sequence is human NKp30 receptor

XX SQ Sequence 190 AA;

Query Match 89.4%; Score 76; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 0.00033;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPPEIRTLGSS 14

|||||

Db 20 WVSQPPPEIRTLGSS 33

RESULT 7

RAY06403

ID AAY06403 standard; protein; 201 AA.

XX AC AAY06403;

XX DT 20-SEP-1999 (first entry)

XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.

XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
 KW signal transduction; immunomodulator; antiinflammatory;
 KW autoimmune disease; inflammation; gene therapy; diagnosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..12 /note= "leader peptide"

FT Protein 13..201 /note= "mature protein"

FT Modified-site 42 /note= "N-glycosylated"

FT Modified-site 68 /note= "N-glycosylated"

FT Modified-site 121 /note= "N-glycosylated"

FT Domain 139..162 /note= "transmembrane domain"

FT Peptide 166..201 /note= "alternatively spliced C-terminal end"

XX WO9923867-A2.

PD 20-MAY-1999.
 XX
 PF 05-NOV-1998; 98WO-US023826.
 XX
 PR 07-NOV-1997; 97US-0064761P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Browning J;
 XX
 XX WPI; 1999-418423/35.
 DR N-PSDB; AAXS9349.
 XX
 XX Novel B-cell myelin oligodendrocyte glycoproteins.
 XX
 XX Claim 2; Page 43; 43pp; English.
 XX
 XX This sequence represents human BMOG, a novel member of the B cell myelin
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre
 CC B cells. 3 C-terminal splice variants (see AY06401-03) of BMOG were
 CC identified. The protein is present primarily in the spleen, in lymph
 CC nodes and in germinal centre B cells. It may have immunoregulatory
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
 CC producing BMOG using these transformed host cells are also provided. BMOG
 CC polypeptides can be used for modulating the immune system of a subject or
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it
 CC with a soluble BMOG protein. The nucleic acid can be used for gene
 CC therapy. The protein can also be used to target a toxin, imaging agent or
 CC radionuclide to a cell expressing BMOG. (All claimed)
 XX
 XX Sequence 201 AA;
 SQ

Query Match 89.4%; Score 76; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 0.00035;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPEIRTLGGS 14
 |||||
 Db 20 WVSQPPEIRTLGGS 33

RESULT 8
 AAE19110
 ID AAE19110 standard; protein; 369 AA.
 XX
 AC AAE19110;
 XX
 DT 29-AUG-2003 (revised)
 DT 21-MAY-2002 (first entry)
 XX
 DE Human NKp30-IgG fusion protein.
 XX
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;
 KW immunoglobulin G; fusion protein.
 XX
 OS Homo sapiens.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FH Region 1..135
 FT /note= "Human NKp30"
 FT Region 136..369
 FT /note= "Human IgG"
 XX
 XX WO200208287-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 19-JUL-2001; 2001WO-IL000664.
 PF

XX 20-JUL-2000; 2000IL-00137419.
 PR
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 PA (UYNE) UNIV BEN-GURION NEGEV.
 XX
 XX Mandelboim O, Forgador A;
 PI
 XX WPI; 2002-195870/25.
 DR N-PSDB; AAD30467.
 XX
 XX New targeting complex capable of targeting an active substance to a
 PT target cell, comprising a target recognition segment and an active
 PT segment, useful for treating pathologies associated with viral infections
 PT or cancer.
 XX
 XX Example 1; Page 108-110; 113pp; English.
 PS
 XX The invention relates to compositions and methods for the treatment and
 CC detection of a variety of viral infections, by using complex agents
 CC comprising the natural killer (NK) cells activating proteins, NKp46 and
 CC NKp44 and functional fragments thereof, linked to therapeutic or imaging
 CC agents. The complex is useful for treating pathologies associated with
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein
 CC -Barr virus, cytomegalovirus, vaccinia virus, EMCV, MVM or herpes virus)
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for
 CC the imaging and monitoring of cancer. The complex may also be used to
 CC detect the presence of abnormal cells in a sample. The antibodies can be
 CC used to qualitatively or quantitatively detect the ligand for the
 CC complex. The present sequence is human NKp30- immunoglobulin G (IgG) FC
 CC region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)
 XX
 XX Sequence 369 AA;
 SQ

Query Match 89.4%; Score 76; DB 5; Length 369;
 Best Local Similarity 100.0%; Pred. No. 0.00065;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPEIRTLGGS 14
 |||||
 Db 20 WVSQPPEIRTLGGS 33

RESULT 9
 ABB70840
 ID ABB70840 standard; protein; 1289 AA.
 XX
 AC ABB70840;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 39312.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX N-PSDB; ABL14943.
 DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 39312; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511); expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB557737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1289 AA;
SQ
Query Match 57.6%; Score 49; DB 4; Length 1289;
Best Local Similarity 81.8%; Pred. No. 57;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 4 QPPEIRTELEGS 14
||| |||||
Db 668 QPPEIRTELEGS 678
RESULT 10
AAU27977
ID AAU27977 standard; protein; 164 AA.
XX
XX AAU27977;
AC
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Human contig polypeptide sequence #130.
DE
XX
XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cystostatic; antirheumatic; antiarthritic; vulnery; antinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX WO200164834-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004926.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX 17-JUN-2000; 2000US-00597707.
PR
XX 14-JUL-2000; 2000US-00616807.
PR
XX 19-SEP-2000; 2000US-00664641.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F, Xu C;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
PI
XX WPI: 2001-589862/56.
DR
XX N-PDDB; AAG44877.
DR

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX
XX Claim 10; Page 141-142; 153pp; English.
XX
XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
CC polypeptides of the invention. The proteins and their associated DNA
CC sequences are useful for the treatment, diagnosis and prevention of
CC various types of disorder in a mammalian subject such as a human, dog,
CC monkey, mouse, hamster or rat. The disorders include cancers such as
CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 164 AA;
SQ
Query Match 52.9%; Score 45; DB 4; Length 164;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 9; Conservative 4; Mismatches 1; Indels 4; Gaps 1;
QY 1 WVSQPP-----EIRTELEGS 14
||| |||||
Db 4 WVAQPPAVTADLQFLEGS 21
RESULT 11
AAG43485
ID AAG43485 standard; protein; 301 AA.
XX
XX AAG43485;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54358.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KW
XX Arabidopsis thaliana.
OS
XX
XX EF1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-00301439.
PF
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX 05-MAR-1999; 99US-0123180P.
PR
XX 09-MAR-1999; 99US-0123548P.
PR
XX 23-MAR-1999; 99US-0125788P.
PR
XX 25-MAR-1999; 99US-0126264P.
PR
XX 29-MAR-1999; 99US-0126785P.
PR
XX 01-APR-1999; 99US-0127462P.
PR
XX 06-APR-1999; 99US-0128234P.
PR
XX 08-APR-1999; 99US-0128714P.
PR
XX 16-APR-1999; 99US-0129845P.
PR


```
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160881P.
PR 22-OCT-1999; 99US-0160889P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161322P.
PR 28-OCT-1999; 99US-0161393P.
PR 28-OCT-1999; 99US-01622142P.

Query Match 52.9%; Score 45; DB 3; Length 301;
Best Local Similarity 72.7%; Pred. No. 57;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPPFIRTL 11
Db 100 WVSQPPKIREL 110

RESULT 12
AAG43484
ID AAG43484 standard; protein; 353 AA.
XX
AC AAG43484;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54357.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0126785P.
PR 06-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136332P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
```

```
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 20-SEP-1999; 99US-0154032P.
PR 22-SEP-1999; 99US-0154779P.
PR 23-SEP-1999; 99US-0155139P.
PR 24-SEP-1999; 99US-0155486P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 21-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.

PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161434P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.9%; Score 45; DB 3; Length 353;
Best Local Similarity 72.7%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSOPPEIRTL 11
Db 152 WSPQPKIREL 162

RESULT 13
ABB92033
ID ABB92033 standard; protein; 353 AA.
AC ABB92033;
XX DT
XX 31-MAY-2002 (first entry)
XX Herbicidally active polypeptide SEQ ID NO 1244.
XX Herbicidal; plant; agriculture; herbicide.
XX Arabidopsis thaliana.
XX WC200210210-A2.
XX PD
XX 07-FEB-2002.
XX PF
XX 28-AUG-2001; 2001WO-EP009892.
XX PR
XX 28-AUG-2001; 2001WO-EP009892.
XX PA (FARB ) BAYER AG.
XX PI Tietjen X, Weidler M;
XX WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX Claim 5; SEQ ID NO 1244; 261pp + Sequence Listing; English.
XX
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX
XX Sequence 353 AA;
```


Query Match 52.9%; Score 45; DB 5; Length 353;
Best Local Similarity 72.7%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRL 11
Db 152 WVSQPPKIREL 162

RESULT 14
AAG43483
ID AAG43483 standard; protein; 380 AA.
AC AAG43483;
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54356.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130443P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140921P.
PR 30-JUN-1999; 99US-0141267P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149336P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149436P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150556P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151085P.
PR 27-AUG-1999; 99US-0151086P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 13-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.9%; Score 45; DB 3; Length 380;
Best Local Similarity 72.7%; Pred. No. 73;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTL 11
| | | | |

Db 179 WSPQPKIREL 189

RESULT 15

ABG23939
ID ABG23939 standard; protein; 185 AA.
XX AC ABG23939;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #23930.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX WO200175067-A2.
XX PN 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR N-PSDB; AAS88126.
XX WPI; 2001-639362/73.
XX DR N-PSDB; AAS88126.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
XX PS Claim 20; SEQ ID NO 54298; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have application in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
amino acid sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 185 AA;

Query Match 51.8%; Score 44; DB 4; Length 185;
Best Local Similarity 75.0%; Pred. No. 50;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEI 8
| | | | |

Db 158 WISQPPEL 165

Search completed: February 26, 2004, 12:10:18
Job time : 12.5103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 5.2122 Seconds
(without alignments)
607.670 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPIRTLEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09A_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	85	100.0	15	13	US-10-036-444-7
2	76	89.4	120	13	US-10-036-444-7
3	76	89.4	190	13	US-10-036-444-7
4	48	56.5	1517	14	US-10-156-761-14713
5	45.5	53.5	83	14	US-10-029-386-32542
6	43	50.6	514	15	US-10-014-099P-59
7	42	49.4	451	9	US-09-902-772-2
8	42	49.4	462	14	US-10-205-823-99
9	42	49.4	462	14	US-10-021-660-95
10	42	49.4	478	9	US-09-902-772-4
11	42	49.4	1317	9	US-09-815-242-5118
12	41.5	48.8	116	15	US-10-264-049-3423
13	41	48.2	87	11	US-09-864-408A-5014
14	41	48.2	137	10	US-09-947-063-2
15	41	48.2	137	10	US-09-947-063-5

16	41	48.2	339	15	US-10-369-493-16318
17	41	48.2	368	14	US-10-166-183-6
18	40	47.1	62	9	US-09-864-761-44601
19	40	47.1	84	9	US-09-135-238B-4
20	40	47.1	196	9	US-09-741-669-368
21	40	47.1	196	9	US-09-912-020-376
22	40	47.1	273	15	US-10-369-493-3987
23	40	47.1	341	14	US-10-233-448-6
24	40	47.1	547	15	US-10-080-334-276
25	40	47.1	869	14	US-10-238-075-957
26	39.5	46.5	1598	10	US-09-863-776-59
27	39.5	46.5	1694	10	US-09-863-776-57
28	39.5	46.5	1694	14	US-10-203-708-36
29	39.5	46.5	1700	10	US-09-863-776-24
30	39.5	46.5	1709	9	US-09-870-759-51
31	39.5	46.5	1709	10	US-09-751-708A-51
32	39.5	46.5	1709	10	US-09-863-776-58
33	39.5	46.5	1709	10	US-09-863-776-60
34	39.5	46.5	1709	14	US-10-203-708-35
35	39	45.9	118	9	US-09-764-868-997
36	39	45.9	170	14	US-10-305-555-36
37	39	45.9	173	14	US-10-305-555-30
38	39	45.9	200	9	US-09-908-193-6
39	39	45.9	230	9	US-09-729-674-186
40	39	45.9	307	9	US-09-816-087-2
41	39	45.9	307	14	US-10-266-643-2
42	39	45.9	307	14	US-10-305-555-2
43	39	45.9	321	9	US-09-920-068A-2
44	39	45.9	321	11	US-09-801-944B-268
45	39	45.9	321	14	US-10-188-405-2

ALIGNMENTS

RESULT 1

US-10-036-444-7
Sequence 7, Application US/10036444
Publication No. US20020142445A1
GENERAL INFORMATION:
APPLICANT: INNATE PHARMA S.A.S.
APPLICANT: UNIVERSITA DI GENOVA
TITLE OF INVENTION: "NO. US20020142445A1 triggering receptor involved in natural cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"
TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same
FILE REFERENCE: SEQ-FR-1060
CURRENT APPLICATION NUMBER: US/10/036,444
CURRENT FILING DATE: 2002-01-07
PRIOR APPLICATION NUMBER: 09/440,514
PRIOR FILING DATE: 1999-11-15
PRIOR APPLICATION NUMBER: 09/456,199
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 7
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide derived from natural sequence, useful for antiserum
OTHER INFORMATION: production
US-10-036-444-7

Query Match 100.0%; Score 85; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WVSQPPPIRTLEGSC 15

Db 1 WVSQPPPIRTLEGSC 15

```
RESULT 2
US-10-036-444-4
; Sequence 4, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; FILE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; FILE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-4
Query Match      89.4%; Score 76; DB 13; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEGS 14
Db 2 WVSQPPPEIRTELEGS 15

RESULT 3
US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; FILE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; FILE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-2
Query Match      89.4%; Score 76; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEGS 14
Db 20 WVSQPPPEIRTELEGS 33

RESULT 4
US-10-156-761-14713
; Sequence 14713, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: KUEHN, Ralf
```

```
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14713
; LENGTH: 1517
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14713
Query Match      56.5%; Score 48; DB 14; Length 1517;
Best Local Similarity 61.5%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEG 13
Db 944 WVSQPPPEIRTELEG 956

RESULT 5
US-10-029-386-32542
; Sequence 32542, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32542
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011311.11
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P81408, EVALUE 5.30e-02
US-10-029-386-32542
Query Match      53.5%; Score 45.5; DB 14; Length 83;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Cy 1 WVSQPPPEIRTELEGSC 15
Db 24 W-SRPPELRPWGSC 37

RESULT 6
US-10-014-099F-59
; Sequence 59, Application US/10014099F
; Publication No. US20040003420A1
; GENERAL INFORMATION:
; APPLICANT: KUEHN, Ralf
```

```

; APPLICANT: FELDER, Susanne
; APPLICANT: SCHWENK, Frieder
; APPLICANT: KUETER-LUKS, Birgit
; APPLICANT: FAUST, Nicole
; TITLE OF INVENTION: Modified Recombinase
; CURRENT APPLICATION NUMBER: US/10/014,099F
; CURRENT FILING DATE: 2001-11-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Xisf recombinase
; US-10-014-099F-59

Query Match          50.8%; Score 43; DB 15; Length 514;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      2 VSQPPPEIRTLGSC 14
Db      414 VESPPEVKTILRAS 426

RESULT 7
US-09-902-772-2
; Sequence 2, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001
; CURRENT APPLICATION NUMBER: US/09/902,772
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: protein sequence from C-11 gene
; US-09-902-772-2

Query Match          49.4%; Score 42; DB 9; Length 451;
Best Local Similarity 46.7%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 15
Db      56 WLSQPPARVTIKMEC 70

RESULT 8
US-10-205-823-99
; Sequence 99, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ganavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-99

Query Match          49.4%; Score 42; DB 14; Length 462;
Best Local Similarity 46.7%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 15
Db      63 WLSQPPARVTIKMEC 77

RESULT 9
US-10-021-660-95
; Sequence 95, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-021-660-95

Query Match          49.4%; Score 42; DB 14; Length 462;
Best Local Similarity 46.7%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 15
Db      63 WLSQPPARVTIKMEC 77

RESULT 10
US-09-902-772-4
; Sequence 4, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et

```

```

; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; TITLE OF INVENTION: the Proteins
; FILE REFERENCE: chugai sei-yaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/09/902,772
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: protein sequence from c-erg gene
US-09-902-772-4

```

```

Query Match      49.4%; Score 42; DB 9; Length 478;
Best Local Similarity 46.7%; Pred. No. 2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY      1 WVSQPPPEIRTLGSC 15
      |::|::|::|::|::|
DB      56 WLSQPPPARVTKMEC 70

```

RESULT 11

```

; Sequence 5118, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chlesen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5118
; LENGTH: 1317
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5118

```

```

Query Match      49.4%; Score 42; DB 9; Length 1317;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 WVSQPPPEIRTLGSC 14
      |::|::|::|::|::|
DB      155 WAAVPAELQTEGS 168

```

RESULT 12

```

; Sequence 3423, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent in Ver. 3.1
; SEQ ID NO 3423
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (90)_FEATURE
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (97)_FEATURE
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3423

```

```

Query Match      48.8%; Score 41.5; DB 15; Length 116;
Best Local Similarity 38.9%; Pred. No. 59;
Matches 7; Conservative 5; Mismatches 3; Indels 3; Gaps 1;

```

```

QY      1 WVSQ---PPEIRTLGSC 15
      |::|::|::|::|::|
DB      10 WLRKLVPTVRSREGTC 27

```

RESULT 13

```

; Sequence 5014, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5014
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-5014

```

```

Query Match      48.2%; Score 41; DB 11; Length 87;
Best Local Similarity 40.0%; Pred. No. 53;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

QY      1 WVSQPPPEIRTLGSC 15

```

Db 68 WVRGPEVKSENCC 82
|| : ||: |
RESULT 14
US-09-947-063-2
; Sequence 2, Application US/09947063
; Publication No. US20030059775A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru et al.
; TITLE OF INVENTION: No. US20030059775A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-112
; CURRENT APPLICATION NUMBER: US/09/947,063
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/229,990
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/229,988
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-063-2

Query Match 48.2%; Score 41; DB 10; Length 137;
Best Local Similarity 46.2%; Pred. No. 83;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 WVSQPPPEIRTLLEG 13
|| : ||: |
Db 105 WVGKDPQVRIVLG 117

RESULT 15
US-09-947-063-5
; Sequence 5, Application US/09947063
; Publication No. US20030059775A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru et al.
; TITLE OF INVENTION: No. US20030059775A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-112
; CURRENT APPLICATION NUMBER: US/09/947,063
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/229,990
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/229,988
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-063-5

Query Match 48.2%; Score 41; DB 10; Length 137;
Best Local Similarity 46.2%; Pred. No. 83;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 WVSQPPPEIRTLLEG 13
|| : ||: |
Db 105 WVGKDPQVRIVLG 117

Search completed: February 26, 2004, 12:33:59
Job time : 7.2122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 2.90451 Seconds
(without alignments)
266.616 Million cell updates/sec

Title: US-10-036-444-7

Sequence: 1 WVSQPEIRTEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/2/iaa/5A-COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B-COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A-COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B-COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/6C-COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/6D-COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	49.4	451	3	US-08-878-177-2
2	42	49.4	478	3	US-08-878-177-4
3	42	49.4	1128	4	US-09-252-991A-31032
4	41	48.2	249	3	US-09-189-760-6
5	41	48.2	249	3	US-09-188-811-6
6	41	48.2	249	3	US-09-514-422-6
7	41	48.2	289	4	US-09-673-395A-155
8	41	48.2	478	4	US-09-134-000C-6219
9	41	48.2	517	3	US-09-189-760-2
10	41	48.2	517	3	US-09-514-422-2
11	41	48.2	567	3	US-09-188-811-2
12	40	47.1	202	4	US-09-328-352-5586
13	40	47.1	344	4	US-09-252-991A-31357
14	39	45.9	146	4	US-09-252-991A-32438
15	39	45.9	650	4	US-09-489-039A-7678
16	38	44.7	353	4	US-09-252-991A-19842
17	38	44.7	409	1	US-08-190-802A-51
18	38	44.7	409	2	US-08-283-917-3
19	38	44.7	409	3	US-08-961-716-3
20	38	44.7	409	3	US-08-477-346-51
21	38	44.7	409	4	US-08-473-089-51
22	38	44.7	409	4	US-08-487-072A-51
23	38	44.7	410	2	US-08-283-917-9
24	38	44.7	410	2	US-08-961-716-9
25	38	44.7	414	4	US-09-543-681A-7732
26	38	44.7	414	4	US-09-543-681A-7982
27	38	44.7	675	3	US-09-171-878-1

Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 387, App
Sequence 387, App
Sequence 387, App
Sequence 27, Appli
Sequence 9057, Ap
Sequence 13, Appli
Sequence 20783, A
Sequence 8738, Ap
Sequence 16, Appli
Sequence 10, Appli
Sequence 2, Appli
Sequence 19546, A
Sequence 6, Appli

28 38 44.7 947 2 US-08-887-518-2
29 38 44.7 947 2 US-09-023-321-2
30 38 44.7 947 2 US-09-032-475-2
31 38 44.7 947 3 US-09-257-703-1
32 38 44.7 947 4 US-09-871-889A-1
33 38 44.7 2972 4 US-08-469-260A-387
34 38 44.7 2972 4 US-08-488-446-387
35 38 44.7 2972 4 US-08-467-344A-387
36 37 43.5 107 4 US-09-345-236B-27
37 37 43.5 256 4 US-09-483-039A-9057
38 37 43.5 263 5 PCT-US91-08177-13
39 37 43.5 280 4 US-09-252-991A-20783
40 37 43.5 370 4 US-09-489-039A-8738
41 37 43.5 449 2 US-08-819-458A-16
42 37 43.5 490 4 US-09-056-285A-10
43 37 43.5 496 2 US-08-463-418-2
44 37 43.5 577 4 US-09-252-991A-19546
45 37 43.5 602 2 US-08-413-652-6

ALIGNMENTS

RESULT 1

US-08-878-177-2
; Sequence 2, Application US/08878177
; Patent No. 6294354
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/08/878,177
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: protein sequence from C-11 gene
US-08-878-177-2

Query Match 49.4%; Score 42; DB 3; Length 451;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WVSQPEIRTEGSC 15

Db 56 WLSQPPARTVTKMEC 70

RESULT 2

US-08-878-177-4
; Sequence 4, Application US/08878177
; Patent No. 6294354
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: chugai seiyaku kabushiki kaisha 5001
; CURRENT APPLICATION NUMBER: US/08/878,177
; CURRENT FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: protein sequence from c-erg gene
US-08-878-177-4

Query Match 49.4%; Score 42; DB 3; Length 478;
Best Local Similarity 46.7%; Pred. No. 70;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
OY 1 WVSQPEIRTEGSC 15
|:|:|:|:|:|:|
Db 56 WLSQPPARVIRKMEC 70

RESULT 3
US-09-252-991A-31032
; Sequence 31032, Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31032
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31032

Query Match 49.4%; Score 42; DB 4; Length 1128;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPEIRTEGSC 14
|:|:|:|:|:|:|
Db 155 WAAVPAELQTOEGS 168

RESULT 4
US-09-189-760-6
; Sequence 6, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-189-760-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPEIRTEGSC 14
|:|:|:|:|:|:|
Db 181 WIETPPSIKSLDSN 194

RESULT 5
US-09-188-811-6
```

```
; Sequence 6, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-188-811-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPEIRTEGSC 14
|:|:|:|:|:|:|
Db 181 WIETPPSIKSLDSN 194

RESULT 6
US-09-514-422-6
; Sequence 6, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-422-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPEIRTEGSC 14
|:|:~|:~|:~|:~|:~|
Db 181 WIETPPSIKSLDSN 194

RESULT 7
US-09-673-395A-155
; Sequence 155, Application US/09673395A
; Patent No. 8620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
```

```
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 155
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-155

Query Match      48.2%; Score 41; DB 4; Length 289;
Best Local Similarity 46.7%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      1 WVSQPPPEIRTLGSC 15
Db      267 WQOPPAARSCYGLC 281

RESULT 8
US-09-134-000C-6219
; Sequence 6219, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6219
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6219

Query Match      48.2%; Score 41; DB 4; Length 478;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 WVSQPPPEIRTLGSC 15
Db      290 WIEMPLLFLVGTG 304

RESULT 9
US-09-189-760-2
; Sequence 2, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-189-760-2

Query Match      48.2%; Score 41; DB 3; Length 517;
Best Local Similarity 35.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 WVSQPPPEIRTLGSC 14
Db      449 WIETPPSIKSLDSN 462

RESULT 10
US-09-514-422-2
; Sequence 2, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-422-2

Query Match      48.2%; Score 41; DB 3; Length 517;
Best Local Similarity 35.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 WVSQPPPEIRTLGSC 14
Db      449 WIETPPSIKSLDSN 462

RESULT 11
US-09-188-811-2
; Sequence 2, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-188-811-2

Query Match      48.2%; Score 41; DB 3; Length 567;
Best Local Similarity 35.7%; Pred. No. 1.2e+02;
```

